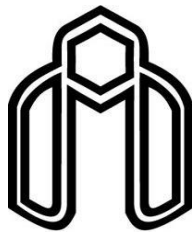


صلى الله عليه وسلم



دانشگاه صنعتی شاهرود

دانشکده مهندسی صنایع و مدیریت

پایان نامه کارشناسی ارشد مدیریت کسب و کار

پایان نامه کارشناسی ارشد

پیش‌بینی شاخص بورس با استفاده از الگوریتم تکامل دستوری

نگارنده: محمدعلی یاراحمدی

استاد راهنما

دکتر سید مجتبی میرلوحی

استاد مشاور

دکتر علی اکبر حسنی

بهمن ۱۳۹۵

شماره: ۳۹۵۵۳۴۷
تاریخ: ۹۶/۱۱/۱۹
ویرایش:

شماره:
تاریخ:
ویرایش:



مدیریت تحصیلات تکمیلی

باسمه تعالی

فرم شماره ۶: صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای محمدعلی یاراحمدی به شماره دانشجویی ۹۳۱۸۳۸۴ رشته MBA گرایش تحت عنوان پیش بینی شاخص کل بورس با استفاده از الگوریتم تکامل دستوری با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می-گردد:

<input type="checkbox"/> مردود	<input type="checkbox"/> دفاع مجدد	<input checked="" type="checkbox"/> قبول (با درجه: <u>بسیار خوب</u> امتیاز: ۱۸/۱۵)
		نوع تحقیق: <input checked="" type="checkbox"/> نظری <input type="checkbox"/> عملی

- ۱- عالی (۲۰ - ۱۹)
۲- بسیار خوب (۱۸/۹۹ - ۱۸)
۳- خوب (۱۶ - ۱۷/۹۹)
۴- قابل قبول (۱۵/۹۹ - ۱۴)
۵- نمره کمتر از ۱۴ غیر قابل قبول

امضاء	مرتبه علمی	نام و نام خانوادگی	عضو هیأت داوران
			۱- استاداراهنمای اول دکتر سید مجتبی میرلوحی
			۲- استاداراهنمای دوم
			۳- استاد مشاور دکتر علی اکبر حسینی
			۴- نماینده شورای تحصیلات تکمیلی دکتر حکمی نسب
			۵- استاد ممتحن اول دکتر رضا شیخ
			۶- استاد ممتحن دوم دکتر محمدعلی مولایی

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء و مهر دانشکده:

دانشگاه صنعتی شاهرود

تقدیم به سه وجود مقدس:

آمان که ناتوان شدند تا ما به توانایی برسیم ...

موباشان سپید شد تا ما رو سفید شویم ...

و عاشقانه سوختند تا گرما بخش وجود ما و رو همسگر راهمان باشند ...

پدرم

مادرم

استادانمان

سپاس‌گذاری

شکر شایان‌نثار ایزدمنان که توفیق را رفیق راهم ساخت تا این پایان نامه را به پایان

برسانم. از استادان فاضل و اندیشمند جناب آقایان دکتر شیخ، دکتر مولایی و جناب

آقای دکتر میرلوحی به عنوان استاد راهنما و دکتر حسینی به عنوان استاد مشاور که همواره

بگذارنده را مورد لطف و محبت خود قرار داده‌اند، کمال تشکر را دارم.

تعهد نامه

اینجانب محمدعلی یاراحمدی دانشجوی دوره کارشناسی ارشد رشته مدیریت MBA دانشکده مهندسی صنایع و مدیریت دانشگاه صنعتی شاهرود نویسنده پایان نامه پیش‌بینی شاخص کل بورس و اوراق بهادار با استفاده از الگوریتم تکامل دستوری تحت راهنمایی دکتر سید مجتبی میرلوحی متعهد می‌شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
 - در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
 - مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
 - کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
 - حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می‌گردد.
 - در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
 - در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است.
- تاریخ امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزار ها و تجهیزیات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

پژوهش های زیادی جهت پیش بینی قابل قبول و قابل اطمینان در حوزه پیش بینی شاخص بورس اوراق بهادار صورت گرفته است. این موضوع از آنجا مد نظر قرار گرفته است که دستیابی به روش ها و فنون جدید پیش بینی دقیق تر نوسان های شاخص ها در بورس از جمله آرمان های سرمایه گذاران و دست اندر کاران آن در سراسر دنیا است. قلمرو زمانی پژوهش از سال ۱۳۸۷ تا آذر سال ۱۳۹۵ و قلمرو مکانی بورس اوراق بهادار تهران است. این پژوهش سعی بر آن دارد که با استفاده از الگوریتم تکامل دستوری شاخص بورس اوراق بهادار را پیش بینی کند. سابق پتانسیل الگوریتم تکامل دستوری برای کشف قوانین مفید جهت پیش بینی صرفاً در حیطه ورشکستگی شرکت ها بکار گرفته شده و متمر ثمر واقع شده است. در اینجا از شاخص کل و شاخص های مرتبط دیگر، که بصورت روزانه و طی دوره ۸ سال اخیر که ثبت شده اند، بکار گرفته شده است. یافته ها نشان می دهد که مدل استفاده شده توانایی قابل قبولی در پیش بینی شاخص کل دارد.

واژه های کلیدی:

مدل های پیش بینی شاخص کل، الگوریتم تکامل دستوری، پیش بینی شاخص کل

فهرست

۱- فصل اول	۱
۱-۱- مقدمه	۲
۲-۱- بیان مسئله	۴
۳-۱- استفاده کنندگان از نتایج تحقیق	۷
۴-۱- نوآوری پژوهش	۷
۵-۱- هدف تحقیق	۷
۶-۱- قلمرو تحقیق	۸
۱-۶-۱- قلمرو موضوعی	۸
۲-۶-۱- قلمرو مکانی	۸
۳-۶-۱- قلمرو زمانی	۸
۲- فصل دوم	۹
۱-۲- مقدمه	۱۰
۲-۲- پیش‌بینی	۱۲
۳-۲- بورس و شاخص‌های آن	۱۳
۴-۲- بازار کارا و امکان پیش‌بینی قیمت‌ها:	۱۷
۵-۲- اشکال مختلف کارایی بازار	۱۹
۱-۵-۲- شکل ضعیف کارایی:	۱۹
۲-۵-۲- شکل نیمه قوی کارایی	۱۹
۳-۵-۲- شکل قوی کارایی:	۲۰
۶-۲- امکان پیش‌بینی قیمت‌ها در بازار:	۲۰
۷-۲- انواع روش‌های پیش‌بینی	۲۵
۱-۷-۲- روش‌های کلاسیک پیش‌بینی	۲۶
۲-۷-۲- رویکرد تجزیه و تحلیل بنیادین:	۲۶
۳-۷-۲- رویکرد تجزیه و تحلیل تکنیکال	۲۷
۴-۷-۲- رویکرد مبتنی بر نظریه‌های مدرن مالی:	۲۸

۲۹	۸-۲- مدل قیمت گذاری دارایی های سرمایه ایی (CAPM).....
۳۳	۹-۲- مدل فاما و فرنچ:.....
۳۴	۲-۹-۱- عامل اندازه شرکت ها:.....
۳۴	۲-۹-۲- عامل ارزش شرکت ها:.....
۳۶	۲-۹-۳- سایر مدل های چند عاملی:.....
۴۰	۲-۱۰-۱- روش های مدرن پیش بینی.....
۴۱	۲-۱۰-۱- رویکرد بی نظمی و پویایی غیر خطی.....
۴۲	۲-۱۱-۱- پیشینه ی پژوهش.....
۴۳	۲-۱۱-۱- پیشینه داخلی.....
۵۰	۲-۱۱-۲- پیشینه خارجی.....
۵۳	۳- فصل سوم.....
۵۴	۳-۱- مقدمه.....
۵۷	۳-۲- انواع الگوریتم های تکاملی.....
۵۷	۳-۳- استراتژی های تکاملی.....
۵۸	۳-۴- برنامه ریزی تکاملی.....
۶۰	۳-۵- الگوریتم ژنتیک.....
۶۱	۳-۶- الگوریتم تکامل دستوری.....
۶۲	۳-۷- گرامر.....
۶۴	۳-۷-۱- رمزگشایی.....
۶۴	۳-۸- تطابق GE و سیستم بیولوژیک.....
۶۷	۳-۹- نگاشت ژنو تیپ به فنو تیپ.....
۶۸	۳-۱۰- تلفیق در GE.....
۶۹	۳-۱۱- جهش در GE.....
۶۹	۳-۱۲- ساختار الگوریتم تکامل دستوری.....
۷۰	۳-۱۲-۱- کروموزوم.....
۷۰	۳-۱۲-۲- جمعیت.....

۷۰	۳-۱۲-۳- تابع برازندگی
۷۰	۳-۱۲-۴- عملگرهای الگوریتم تکامل دستوری
۷۵	۳-۱۲-۵- مزیت های الگوریتم تکامل دستوری نسبت به الگوریتم ژنتیک
۷۶	۳-۱۲-۶- روند کلی الگوریتم تکامل دستوری
۷۸	۳-۱۲-۷- روند کلی بهینه سازی و حل مسائل در الگوریتم تکامل دستوری:
۷۹	۳-۳- مشخص شدن مکان فرزند تولید شده در کروموزوم
۷۹	۳-۴- پذیرش: جا دادن فرزند جدید در داخل جمعیت.
۷۹	۳-۱۲-۸- شرط پایان الگوریتم.
۸۱	۳-۱۳- نتیجه گیری
۸۳	۴- فصل چهارم
۸۴	۴-۱- مقدمه
۸۴	۴-۲- تعریف مسئله
۸۴	۴-۳- انتخاب متغیرها
۸۵	۴-۴- تحلیل داده ها
۸۵	۴-۵- آماده سازی داده ها
۸۶	۴-۶- مدل سازی
۸۶	۴-۶-۱- کد نویسی
۸۹	۴-۷- نتایج به دست آمده از الگوریتم تکامل دستوری در نرم افزار متلب
۸۹	۴-۷-۱- نتایج حاصل برای یک روز آتی
۸۹	۴-۷-۲- نتایج حاصل برای روز دوم
۹۰	۴-۷-۳- نتایج حاصل برای روز سوم
۹۰	۴-۸- نقطه تمایز افزایش یا کاهش شاخص در جدول زیر ارائه می شود.
۹۱	۴-۹- شاخص های اثر گذار برای هر روز در جدول زیر آمده است.
۹۱	۴-۹-۱- مدل ارائه شده برای پیش بینی یک روز آتی
۹۱	۴-۹-۲- مدل ارائه شده برای پیش بینی روز دوم آتی

- ۹۱..... ۴-۹-۳- مدل ارائه شده برای پیش‌بینی روز سوم آتی
- ۹۲..... ۴-۱۰- اعتبار سنجی
- ۹۷..... ۵- فصل پنجم:
- ۹۸..... ۵-۱- نتایج حاصل از مدل GE در پیش‌بینی شاخص کل
- ۹۸..... ۵-۲- بحث و نتیجه‌گیری
- ۹۹..... ۵-۳- محدودیت‌ها
- ۱۰۰..... ۵-۴- پیشنهادات
- ۱۰۱..... منابع و مآخذ
- ۱۰۵..... پیوست

فهرست اشکال

- شکل ۱-۲: نمونه ای تحلیل قیمت سهام با ابزارهای رویکرد تحلیل تکنیکال..... ۲۸
- شکل ۱-۳: طرح کلی الگوریتم تکاملی..... ۵۶
- شکل ۲-۳: شبه کد برنامه ریزی تکاملی..... ۶۰
- شکل ۳-۳: یک نمونه از گرامر در فرم BNF..... ۶۳
- شکل ۴-۳: تطابق سیستم GE و سیستم بیولوژیکی تولید پروتئین..... ۶۵
- شکل ۵-۳: فرایند نگاشت GE..... ۶۶
- شکل ۶-۳: یک گرامر علامت گذاری شده که قوانین تولید با شاخص های مرتبط در آن نشان می دهد..... ۶۸
- شکل ۷-۳: نحوه ارزیابی شایستگی در چرخ رولت..... ۷۲
- شکل ۸-۳: یک نمونه از تلفیق (آمیزش)..... ۷۳
- شکل ۹-۳: ادغام دو نقطه ای..... ۷۴
- شکل ۱۰-۳: تلفیق جامع..... ۷۴
- شکل ۱۱-۳: یک کروموزوم قبل و بعد از اعمال..... ۷۵
- شکل ۱۲-۳: شماتیک ساده و کلی از الگوریتم تکامل دستوری..... ۷۷
- شکل ۱۳-۳: نحوه ارزیابی تابع شایستگی در چرخ رولت (براساس الگوریتم ژنتیک)..... ۷۸

فهرست جداول

- جدول ۳-۱: قبل از تلفیق..... ۸۱
- جدول ۳-۲: عمل بعد از تلفیق..... ۸۱
- جدول ۴-۱: متغیرهای حذف شده حین اجرای برنامه..... ۸۷
- جدول ۴-۲: نتایج پیش‌بینی برای یک روز آینده همراه با متغیرهای مؤثر..... ۸۹
- جدول ۴-۳: نتایج پیش‌بینی برای روز دوم همراه با متغیرهای مؤثر..... ۸۹
- جدول ۴-۴: نتایج پیش‌بینی برای روز سوم همراه با متغیرهای مؤثر..... ۹۰
- جدول ۴-۵: شرایط افزایش یا کاهش شاخص..... ۹۰
- جدول ۴-۶: شاخص‌های اثرگذار برای پیش‌بینی هر روز..... ۹۱
- جدول ۴-۷: میانگین دوره در نظر گرفته شده برای شاخص‌های مؤثر..... ۹۲
- جدول ۴-۸: مقادیر محاسبه شاخص برای روز اول..... ۹۳
- جدول ۴-۹: مقادیر محاسبه شاخص برای روز دوم..... ۹۳
- جدول ۴-۱۰: مقادیر محاسبه شاخص برای روز سوم..... ۹۴
- جدول ۴-۱۱: اعتبار سنجی پیش‌بینی با واقعیت (دی ماه ۹۵)..... ۹۴
- جدول ۵-۱: متغیرهای مؤثر و ریسک‌زا برای هر یک از روزهای پیش‌بینی شده..... ۹۹

۱- فصل اول

کلیات پژوهش

بشر در دنیای امروزی به صورت روزمره در بازارهای گوناگون درگیر تصمیم گیری های بیشماری بوده و هر گونه پیشنهادی که امکان بهبود دقت و صحت تصمیم و یا کاهش زمان تصمیم گیری را برای او به ارمغان بیاورد برای وی جذاب و ارزشمند می باشد. یکی از بازارهایی که امروزه رو به رونق بوده و مزایای الگوریتم های تکاملی در آن بسیار مشهود می باشد بازارهای پولی و سرمایه شامل بازار بورس اوراق بهادار می باشد. فعالان این بازار به خرید و فروش سهام شرکتها در آن بازار پرداخته و از آن طریق با پذیرفتن ریسک آینده سهم برای خود سود و یا زیان به بار می آورند (نوروزی موحد، ۱۳۹۴). ناشناخته بودن عوامل تاثیر گذار بر تغییرات قیمت سهام همواره دلیلی برای روی آوردن به پیش بینی قیمت سهام شرکت ها است. امروزه مدیران مالی ترجیح می دهند مکانیزمی در اختیار داشته باشند که بتواند آن ها را در امور تصمیم گیریشان یاری نماید به همین دلیل توجه به روش های پیش بینی بسیار مورد توجه قرار گرفته است. سرمایه گذارانی که در بورس اوراق بهادار سرمایه گذاری می کنند برای آنکه بتوانند سود دریافتی خود را حداکثر کنند نیازمند پیش بینی ارزش آتی سهام هستند (بیگی ، مقایسه رویکرد تکنیکی و بنیادی در پیش بینی قیمت سهام و ارائه یک مدل ترکیبی با کمک سیستم های هوشمند، ۱۳۹۰). پیش بینی آینده در عرصه اقتصاد و بازار سرمایه یکی از مهمترین مسائل مورد بحث در علوم مالی بوده است. شوک های ناگهانی بازار و سقوط قیمت ها تعداد زیادی از سرمایه گذاران را از بازار خارج می کند (فلاح شمس & دلنواز، ۱۳۸۸). موضوع پیش بینی شاخص کل بورس تاکنون مورد توجه بسیاری از تحلیل گران حوزه مدیریت مالی و حسابداری بوده است. زیرا افزایش این شاخص نشان دهنده رشد و توسعه اوضاع و احوال اقتصادی و کاهش آن بیانگر رکود و بحران است. شاخص یک معیار آماری است که تغییر حرکت و جهت یک اقتصاد یا یک بازار سهام را نشان می دهد. شاخص ها ابزار سودمندی برای ردیابی روند های بازار به خصوص در کوتاه مدت هستند. شاخص های بورس معمولاً به عنوان یکی از معیارهای مهم رونق یا رکود

اقتصاد داخلی نیز بکار می روند. تحلیل گران مالی معتقدند که درک شاخص و تغییرات آن می تواند به سرمایه گذاران در تصمیمات سرمایه گذاری مناسب تر یاری رساند. در بازار های مالی اساساً شاخص، یک پرتفوی فرضی از اوراق بهادار است که شامل کل بازار مربوطه یا بخشی از آن است (مودی و لوین، ۱۹۹۳).^۱ در بورس اوراق بهادار شاخص کل، پرتفویی فرضی از کلیه سهام پذیرفته شده است در حالی که شاخص مالی، در برگیرنده پرتفویی فرضی از شرکت های سرمایه گذاری و هولدینگ است (محمدی تاکامی، ۱۳۹۲). پیش بینی این شاخص می تواند در تصمیم گیری سرمایه گذاران، صاحبان صنایع و حتی تحلیل گران بازار سرمایه و اقتصاد مفید فایده واقع شود. بورس آئینه تمام نمای وضعیت اقتصادی یک کشور است این کار در سالهای اخیر از تکنیک هایی چون شبکه های عصبی که جایگاه ویژه ای در پیش بینی تغییرات و رویدادهای بازار سرمایه دارند استفاده شده است. الگوریتم های تکاملی تکنیک های اجرای مکانیسم هایی هستند که از تکامل بیولوژیکی از قبیل تولید مثل، جهش، با ترکیبی، انتخاب طبیعی و مناسب ترین بقاء الهام گرفته اند (اونیل^۲ و همکاران، ۲۰۱۲). در این تحقیق برای پیش بینی شاخص بورس از یکی از رو های تکاملی به نام تکامل گرامری یا دستوری یا Grammatical Evolution استفاده شده است. تکامل گرامری که می تواند برای تکامل برنامه های کامپیوتری بکار رود، مجموعه ای از دستور زبان یا جملات عمومی در هر زبان است. مجموعه قوانین می تواند متنوع تر از مدل رگرسیون یا یک سیستم تجاری برای بازار مالی باشد.

الگوریتم GE یک شکل از برنامه نویسی ژنتیکی که براساس دستور زبان است و به تعدادی از حوزه های مسئله که در بر گیرنده برنامه های کامپیوتر و پیش بینی مالی است اعمال می شود (باربازون^۳ و همکاران، ۲۰۰۶). GE برای سطوح مختلفی از مسائل دربرگیرنده نمونه های موفق مالی بکار برده شده است. نقطه

¹ J. Moody, U. Levin

² O'Neill

³ Anthony Brabazon

قوت GE استفاده از یک دستور زبان برای ترکیب حوزه دانش، برای مسئله ای که تحقیق پیش رو برای حل آن تلاش می کند، است (اونیل و همکاران، ۲۰۱۲). روش GE دارای مزیت هایی دیگر از قبیل: سهولت در پیاده سازی آن، تولید برنامه به زبان دلخواه و استفاده از آخرین پیشرفت های الگوریتم تکاملی است، در صورتی که ما بقی روش های تکاملی دارای این مزیت نیستند. با توجه به مزیت های ذکر شده در مورد GE استفاده از این الگوریتم در پیش بینی شاخص بورس می تواند موضوعی راهگشا و بدیع تلقی شود.

۱-۲- بیان مسئله

پیش بینی، فرایند برآورد موقعیت های ناشناخته است. یک پیش بینی یک پیشگویی در مورد رویدادهای آینده در اختیار سرمایه گذاران بازار مالی می گذارد و می تواند تجارب گذشته را به پیش بینی حوادث آینده بدل سازد. در سالهای اخیر پیش بینی، به پیش بینی شیوه طرح تقاضا در کسب و کار روزانه شرکت های سازنده تبدیل شده است. پیش بینی با میزان دقت مناسب یک ابزار مهم و قدرتمند در فرایند برنامه ریزی است. معمولاً پیش بینی به دو روش صورت می گیرد. روش اول روش های آماری است که پارامترهای مدل پیش بینی به صورت احتمالی برآورد می شود. روش دوم روش های ریاضی است که پارامترهای مدل به صورت قطعی برآورد می شوند. بنابراین می توان با در نظر گرفتن متغیرهای مختلف برای شاخص بورس افزایش یا کاهش آن را پیش بینی نمود. روشهای پیش بینی دقیق برای مدیریت سبد سهام توسط بانک های سرمایه گذاری و تجاری بسیار اهمیت دارد (تهرانی و مراد پور، ۹۱). از اواسط دهه ۷۰ و به ویژه از سال ۱۹۸۰ تلاش های گسترده ای در زمینه قابلیت پیش بینی قیمت های سهام با استفاده از روش های ریاضی جدید، سری زمانی طولانی و ابزار پیشرفته تری مثل هوش مصنوعی آغاز شد و آزمون های زیادی بر روی اطلاعات قیمت و شاخص سهام در کشورهای مثل انگلستان، آمریکا، کانادا، آلمان و ژاپن صورت گرفت تا وجود یا فقدان ساختاری معین در اطلاعات قیمت سهام نشان داده شود و از

این راه فرضیه گام های تصادفی را نقض کند (سینایی و همکاران، ۸۴). پیش‌بینی قیمت سهام می‌تواند تا حدودی عدم اطمینان موجود در بازدهی سهام را کاهش دهد و یک سرمایه گذار بتواند با کمک آن به سودهای با ثبات تری دست یابد، بنابراین پیش‌بینی قیمت سهام می‌تواند عاملی جهت سوق دادن سرمایه گذاران به سوی بورس و اوراق بهادار باشد و این بازار رونق یابد (بیگی، ۱۳۹۰). وجود سرمایه پویا و قابل اعتماد در یک کشور از الزامات نیل به رشد اقتصادی بالا و بالا بردن میل به سرمایه گذاری می باشد. بازار بورس اوراق بهادار جزء پنج بازار اصلی سرمایه ای کشور است که کوچکترین نوسانات عملیاتی آن اثرات بسیاری در اقتصاد کشور خواهد گذاشت. در این بازار بیش از ۴۵۰ شرکت سرمایه پذیر در قالب ۳۵ صنعت به عرضه سهام پرداخته اند و معامله گران حقوقی و حقیقی بیشماری از محل تجارت سهام شرکتها در این بازار به مدیریت اقتصاد بنگاه ها و خانوار می پردازند. در شرایط ثبات اقتصادی کشورها، سود حاصل از فعالیت در بازار بورس از سود بهره بانکی قطعاً بالاتر بوده و این جذابیت سبب می شود در شرایط ثبات و یا رشد اقتصادی کشور، جریان سرمایه از بسیاری از بازارها و همچنین جریان سرمایه های راکد و یا فعال در بازارهای سیاه به سمت بازار بورس اوراق بهادار باشد. در کنار نگاه اقتصادی به بازار بورس، از نگاه اجتماعی نیز حضور معامله گران در سنین جوانی و یا پایان میان سالی، یعنی دقیقاً در سنینی که هیجان ریسک پذیری، انگیزه فعالیتهای پرچالش و توام با ریسک را به همراه دارد، در این بازار شایان توجه است. از این منظر انتظار آن خواهد بود که در صورت وجود موج اخبار رونق و بازدهی قابل قبول این بازار، نسبت عمده ای از جمعیت جویای کار در این بازار به صورت مستقل به اشتغال درآیند. با این توضیحات، می توان نتیجه گرفت حل مساله این تحقیق با ارائه مدلی بابت پیشگویی قابل اعتماد قیمت سهم شرکتها در بازار بورس دارای نه تنها اهمیت از منظر اقتصادی بلکه در زمره مسائل با اهمیت اجتماعی نیز قرار می گیرد. از منظر علم اقتصاد بورس اوراق بهادار از مزایای فراوانی برخوردار است که مهمترین آنها عبارتنداز:

۱. کمک به گسترش مالکیت عمومی: تقسیم مالکیت های بزرگ از طریق فروش سهام آنها در بورس اوراق بهادار، از دید کلان اقتصادی به هدف های توزیع عادلانه درآمد و نیز احساس مشارکت عموم افراد جامعه در فعالیت های تولیدی و تجاری کمک زیادی می کند. بسیاری جامعه شناسان توزیع مالکیت را بین عموم طبقات به لحاظ اجتماعی مفید می دانند. چرا که از لحاظ دو قطبی شدن جامعه جلوگیری می کند و ناآرامی های اجتماعی را کاهش می دهد.

۲. تخصیص مطلوب منابع: یک بازار رقابت کامل در بورس اوراق بهادار، تخصیص مطلوب منابع را میسر می سازد. بورس اوراق بهادار با تعیین بهای رقابتی سهام بعنوان ابزار انتظامی اقتصاد، مدیران غیر کارا و ضعیف را تنبه و مدیران مبتکر و قدرتمند را تشویق می کند. در چنین شرایطی یک شرکت زیان ده نمی تواند از طریق بورس به تأمین مالی خود اقدام نماید.

۳. تسهیل فرایند تأمین مالی شرکت ها: انتشار سهام و فروش اوراق بهادار یکی از مطمئن ترین راههای جمع آوری وجوه لازم برای سرمایه گذاری است. واحد های تولیدی و تجاری به جای دریافت تسهیلات از نظام بانکی می توانند تحت شرایطی از طریق انتشار سهام یا فروش اوراق بهادار در بورس، نیازهای مالی خود را تأمین کنند.

روش Grammatical Evolution یکی از تکنیک های نسبتاً جدید در حوزه محاسبات تکاملی که به منظور برنامه نویسی خودکار استفاده می شود، تکنیک تکامل گرامری، یا باختصار GE است. GE توسط مایکل اونیل و کونر رایان در سال ۱۹۹۸ معرفی شد. این دو به همراه کولینز در دانشگاه لیمبرگ کشور ایرلند از پیش قراولان این عرصه به شمار می روند. برخلاف GP، GE پروسه تکامل را بر روی برنامه های واقعی انجام نمی دهد، بلکه این کار را روی کروموزوم هایی از رشته های باینری با طول متغیر تشکیل شده اند، انجام می دهد. سپس از یک رویه نگاشت برای ایجاد برنامه به هر زبان دلخواه استفاده می شود. این رویه از رشته های باینری برای انتخاب قوانین تولید استفاده می کند. قوانین نیز به فرم گرامری BNF

هستند. در نهایت برنامه ای حاصل می شود که از نظر نحوی درست بوده و می توان آن را توسط یک تابع برازندگی، ارزیابی کرد (دمپسی & همکاران, ۲۰۰۹).

۱-۳- استفاده کنندگان از نتایج تحقیق

یکی از مهمترین بخش های اقتصادی کشورها، بازار سرمایه که ارتباط نزدیکی با ساختار اقتصادی کشور داشته و قوت و ضعف آن می تواند نشان دهنده وضعیت اقتصادی کشور باشد (میرزا خانی و چوبینه, ۱۳۸۹). با توجه به آنکه بورس بعنوان یکی از مهمترین نهادهای بازار سرمایه در ایران محسوب می شود، بنابراین پیش بینی شاخص بورس با استفاده از روش GE می تواند برای بازار اوراق قرضه، بازار سهام، بازار مشتقات مالی بازار خارج بورس (بازار نقد، پیمان آتی، قرارداد سواپ و اختیار معامله)، بازار تبادل ارز و دیگر بازارهای همانند پول، بیمه، کالا و املاک مورد استفاده قرار می گیرد.

۱-۴- نوآوری پژوهش

همانطور که قبل تر ذکر شد در زمینه پیش بینی شاخص بورس پژوهش های زیادی صورت گرفته است اما با استفاده الگوریتم های تکامل دستوری تا کنون در زمینه پیش بینی شاخص بورس انجام نگرفته است. استفاده از این الگوریتم بسیاری از روندهای غیر قابل فهم در بورس را برای ما تسهیل می کند. پیش بینی شاخص بورس بعنوان یک متغیر کلان اقتصادی و آئینه تمام نمای اقتصاد کشور می توان تصمیمات مناسب وضعیت اقتصاد کشور گرفت.

۱-۵- هدف تحقیق

این تحقیق به دنبال پیش بینی شاخص بورس اوراق بهادار بازار تهران از روش GE می باشد. برای اجرای روند مذکور و نتیجه گیری در حد واقعیت، این تحقیق باید از نرم افزار Mathwork Matlab R2013b و انواع آزمون ها (آزمون ناهمسانی واریانس، آزمون های اثر ثابت و ...) کمک گرفت.

۱-۶-قلمرو تحقیق

۱-۶-۱-قلمرو موضوعی

موضوع تحقیق با توجه به اینکه پیش‌بینی بازار بورس در ایران است بنابراین می‌توان آنرا در حیطه دانش مالی و حسابداری به حساب آورد.

۱-۶-۲-قلمرو مکانی

قلمرو مکانی تحقیق بازار بورس و اوراق بهادار تهران است.

۱-۶-۳-قلمرو زمانی

زمان آغاز این پژوهش آبان ۱۳۹۴ است. همچنین بررسی پژوهش بر روی شرکت های پذیرفته شده در بازار بورس و اوراق بهادار تهران از ابتدای سال ۱۳۸۷ تا پایان سال ۱۳۹۳ است.

۲- فصل دوم

ادبیات و پیشینه تحقیق

۲-۱-مقدمه

پیش‌بینی آینده در عرصه اقتصاد و بازار سرمایه یک بحث جدی در علم مالی محسوب می‌شود. امروزه با توجه به اهمیت و گسترش روز افزون بازار سرمایه و جمع‌آوری سرمایه‌های کوچک فردی به سمت فعالیت‌های تولیدی، شناسایی رفتار سرمایه‌گذاران و متغیرهای تأثیرگذاری بر قیمت بازده سهام در این بازارها اهمیت زیادی پیدا کرده است. از این رو متخصصان بازار سرمایه، سالیان متمادی به مطالعه بازار و شناسایی الگوهای مختلف برای پیش‌بینی پرداخته‌اند که برای این امر ترکیبی از تشخیص الگو و تجربه‌ی مبتنی بر مشاهده روابط علت و معلول را بکار بسته‌اند. همچنین برنامه‌های نرم‌افزاری بسیاری نیز وجود دارند که به این تصمیم‌گیری کمک می‌کند و به عنوان موتور پیش‌بینی مورد استفاده قرار می‌گیرند. با این وجود در روندهای مالی، اغلب شرایطی بوجود می‌آید که قوانین را به هم می‌ریزد و پیش‌بینی را توسط روشهای مذکور دشوار می‌سازد.

در ادبیات موضوع، روش‌های پیش‌بینی گوناگونی وجود دارد. این تکنیک‌ها براساس نوع ابزار و نوع داده‌های مورد استفاده به چهارگروه طبقه‌بندی نمود:

۱- روش‌های تحلیل فنی^۱

۲- روش‌های تحلیل بنیادی^۲

۳- روش‌های پیش‌بینی سری‌های زمانی کلاسیک^۳

۴- روش‌های هوشمند^۴.

¹ Technical Analysis

² Fundamental Analysis

³ Traditional Time Series

⁴ Machine Learning

اهمیت ویژه بازار سرمایه در توسعه اقتصادی از طریق هدایت موثر سرمایه‌ها و تخصیص بهینه منابع غیرقابل انکار است. سرمایه‌گذاری در بازار سرمایه مستلزم تصمیم‌گیری می‌باشد که این خود نیازمند دستیابی به اطلاعاتی در خصوص وضعیت آینده قیمت بازار سهام می‌باشد. لذا در صورتی که بتوان روند آتی بازار سهام را با روش‌های مناسب پیش‌بینی نمود، سرمایه‌گذاری می‌تواند بازده حاصل از سرمایه‌گذاری خود را بیشینه سازد. بدیهی است که سرمایه‌گذاری در بورس، بخش مهمی از اقتصاد کشور را تشکیل می‌دهد و بدون شک بیشترین میزان سرمایه از طریق بازارهای بورس در سرتاسر جهان مبادله می‌شود و اقتصاد ملی به شدت متأثر از عملکرد این بازار می‌باشد. علاوه بر آن، بازار بورس هم برای سرمایه‌گذاران حرفه‌ای و هم برای عموم مردم یک ابزار سرمایه‌گذاری در دسترس محسوب می‌شود. بازارهای بورس از متغیرهای کلان اقتصادی و غیر اقتصادی و متغیرهای بسیاری دیگری تأثیر می‌پذیرند، متعدد بودن عوامل مؤثر بر بازارهای سرمایه و ناشناخته بودن آنها، موجب عدم اطمینان در زمینه سرمایه‌گذاری شده است، بنابراین محققان سعی دارند تا با استفاده از تکنیک‌ها و روش‌های مختلف، فاکتورهای مهم و مؤثر در سرمایه‌گذاری شناسایی و پیش‌بینی کرده و تا آنجا که امکان دارد درصد اطمینان از بازده سرمایه‌گذاری را افزایش دهند. روش‌های کلاسیک مانند رگرسیون اگرچه توفیقات نسبی در این زمینه‌ها داشته‌اند اما نتایج آن نتوانسته است پژوهشگران این عرصه را راضی نماید، معمولاً به منظور پیش‌بینی وقایعی که در آینده اتفاق می‌افتد به اطلاعات به دست آمده از رویدادهای تاریخی اتکا می‌شود (فلاح شمس و دلنواز، ۱۳۸۸). به این ترتیب که داده‌های گذشته مورد تجزیه و تحلیل قرار می‌گیرد تا از آن‌گویی قابل‌تعمیم به آینده حاصل گردد، در اغلب روش‌های پیش‌بینی فرض بر این است که روابط بین متغیرها در آینده نیز ادامه خواهد داشت. بنابراین باید برای تحلیل و نتیجه‌گیری از پژوهش باید به دو مهم توجه نمود، اولاً جمع‌آوری داده و اطلاعات لازم و تأیید صحت و سقم آنها، ثانیاً اجرای فرایند برنامه‌ریزی و تبدیل داده‌ها به ورودی برای بدست آوردن خروجی مناسب.

۲-۲- پیش‌بینی

کارآیی نهایی هر تصمیم، به طبیعت یک دنباله از حوادث بستگی دارد که در پی آن تصمیم رخ می‌دهد. اگر بتوان جنبه‌های غیر قابل کنترل این حوادث را قبل از تصمیم‌گیری حدس زد، امکان تصمیم‌گیری بهتری به وجود می‌آید. بنابراین چون حوادث آینده در فرآیند تصمیم‌گیری نقش عمده‌ای را ایفا می‌کند، پیش‌بینی حوادث حائز اهمیت است و هر تصمیم‌گیری آگاهانه نیاز به پیش‌بینی دارد. تحلیل‌گران فنی سعی می‌کنند بر اساس الگوهای موجود در نمودار داده‌های مربوط به بازار، قیمت را پیش‌بینی کنند. تحلیل‌گران بنیادی، با توجه به ارزش واقعی و ذاتی یک سهم اقدام به پیش‌بینی می‌نمایند. در پیش‌بینی با روش‌های کلاسیک، فرض بر این است که مقادیر آینده قیمت، سیر خطی مقادیر گذشته را می‌پیمایند. روش‌های هوشمند الگوهای خطی و غیر خطی موجود در داده‌های مربوط به بازار را دنبال می‌کنند تا بدین وسیله فرایند ایجاد آنها را حدس بزنند (عباس پور م.، ۱۳۸۱). در تجزیه و تحلیل کمی علمی، غالباً پیش‌بینی به صورت استفاده از اطلاعات حال و گذشته در قالب الگوهای مختلف و بکار بردن الگوی مورد نظر برای دوره‌های بعدی است. بدین جهت می‌توان گفت که پیش‌بینی عبارت از برآورد وقایع آینده براساس اطلاعات حال و گذشته است. به طور کلی، روش‌های پیش‌بینی به دو گروه اصلی روش‌های کمی و روش‌های کیفی تقسیم می‌شوند. برای پیش‌بینی وقایع آینده با روش کیفی، بطور کلی از نظرات و عقاید متخصصین استفاده می‌شود. معمولاً وقتی که داده‌های زمانی مربوط به گذشته به اندازه کافی در دسترس نباشد، از روش‌های پیش‌بینی کیفی استفاده می‌شود. روش‌های پیش‌بینی کیفی را می‌توان بر مبنای ذهنی با بکار بردن قضاوت، درک مستقیم و اطلاعات مناسب انجام داد. روش دلفی، روش مقایسه فنی مستقل زمانی، روش برآورد ذهنی، روش تقابلی و روش تحقیق شکل‌شناسی از روش‌های کیفی پیش‌بینی است. روش‌های کمی پیش‌بینی زمانی بکار میرود که داده‌های مربوط به گذشته در دسترس باشد. مدل کمی پیش‌بینی به دو نوع مدل‌های تک متغیره و مدل علی تقسیم می‌شوند. در مدل تک

متغیره تنها براساس الگوی تاریخی متغیر مورد نظر، ارزش آینده آن را پیش‌بینی می‌کنند. در مدل‌های علی با فرض ادامه داشتن الگوی تاریخی در آینده و با استفاده از روابط میان متغیر مورد نظر و سایر متغیرها، ارزش آتی متغیر مورد نظر پیش‌بینی می‌شود. در مدل‌های علی یک متغیر وابسته و یک یا چند متغیر مستقل وجود دارد. روش‌های کمی زمانی بکار گرفته می‌شوند که انتظار می‌رود الگوی داده‌ها در آینده نیز ادامه داشته باشد و از روش‌های کیفی زمانی استفاده می‌شود که انتظار می‌رود الگوی داده‌ها تغییر کند. پیش‌بینی‌های ناشی از روش‌های کمی، اغلب مورد ارزیابی ذهنی نیز قرار می‌گیرند. این ارزیابی، ممکن است منجر به اصلاح و تعدیل پیش‌بینی گردد. در روش‌های کمی تک متغیره، برای انجام عمل پیش‌بینی، داده‌های سری زمانی مورد استفاده قرار می‌گیرد. در یک تعریف کلی، سری زمانی مجموعه‌ای از مشاهدات است که بر حسب زمان مرتب شده‌اند، به عبارتی دیگر، سری زمانی عبارت است از توالی زمانی مشاهدات مربوط به یک متغیر معین است. در این پژوهش برای یافتن الگویی که در فراهم کردن پیش‌بینی مورد نظر مؤثر واقع شود، داده‌های سری زمانی را مورد بررسی قرار می‌دهیم. تجزیه و تحلیل سری‌های زمانی به داده‌هایی مربوط می‌شود که مستقل نبوده و به طور متوالی به هم وابسته‌اند؛ در تجزیه و تحلیل سری‌های زمانی، وابستگی بین مشاهدات متوالی مورد توجه و استفاده قرار می‌گیرد. یکی از کاربردهای تحلیل سری‌های زمانی پیش‌بینی است که حوزه وسیعی از آمار عملی و اقتصاد کاربردی را تشکیل می‌دهد.

۲-۳- بورس و شاخص‌های آن

در علم اقتصاد، بورس به بازاری اطلاق می‌شود که قیمت‌گذاری و خرید و فروش کالا و اوراق بهادار در آن انجام می‌پذیرد و در یک طبقه بندی کلی شامل بورس کالا و اوراق بهادار است که بازار بورس اوراق بهادار در ایران در سال ۱۳۴۶ راه‌اندازی شده است. امروزه بورس به عنوان ابزاری بسیار مهم از بازار سرمایه، نقش ویژه‌ای را در بازار اقتصادی ایفا می‌کند و با قیمت‌گذاری، کاهش ریسک، تجهیز منابع و

تخصیص بهینه سرمایه، زمینه را برای رونق اقتصادی فراهم می کند (پاکدامن، ۱۳۸۷). بورس اوراق بهادار به معنای یک بازار متشکل و رسمی سرمایه است که در آن خرید و فروش سهام شرکت ها یا اوراق قرضه دولتی یا مؤسسات معتبر خصوصی، تحت ضوابط و قوانین و مقررات خاصی انجام می شود. مشخصه مهم بورس اوراق بهادار، از سویی مرکز جمع آوری پس اندازها و نقدینگی بخش خصوصی به منظور تأمین مالی پروژه های سرمایه گذاری بلند مدت است و از سوی دیگر، مرجع رسمی و مطمئنی است که دارندگان پس اندازهای راكد می توانند محل نسبتاً مناسب و ایمن سرمایه گذاری جستجو کرده و وجوه مازاد خود را برای سرمایه گذاری در شرکت ها به کار انداخته و یا با خرید اوراق قرضه دولت ها و شرکت های معتبر، از سود معین و تضمین شده ای برخوردار شوند (جوادی، ۱۳۶۹). سرمایه گذاران و مسئولان بورس اوراق بهادار به منظور دستیابی به تصویر مناسبی از روند این بازار و توانایی ارزیابی گذشته و در مواردی پیش بینی وضعیت اقتصادی آینده، از شاخص های بورس بهره می برند. تجزیه و تحلیل دقیق تر روند قیمت در بورس اوراق بهادار محتاج شاخص هایی با کارکردهای گوناگون است و بدین سبب امروزه شاخص های بسیار متنوعی در بورس های معتبر جهانی محاسبه و منتشر می شوند (کیانی، ۱۳۹۰). شاخص بورس همچون دماسنج نشان دهنده وضعیت بازار سرمایه و وضعیت اقتصادی یک کشور است. کاهش شاخص در حالت کلی به معنای رکود اقتصادی و افزایش آن به معنای رونق اقتصادی است (دوانی، ۱۳۸۱). به طور خلاصه شاخص وسیله ای برای اندازه گیری و مقایسه پدیده هایی است که دارای ماهیت و خاصیت مشخصی هستند. بنابراین بر مبنای شاخص می توان تغییرات ایجاد شده در متغیرهای معینی را در طول یک دوره بررسی نمود (غفاری، ۱۳۸۳). شاخص قیمت سهام، هم از دید سرمایه گذاران برای سرمایه گذاری در سهام خاص و هم به عنوان یک شاخص اقتصادی از دید اقتصاد کلان در جامعه کاربرد وسیعی دارد. واژه شاخص به معنای نمودار، نماینده، نشان دهنده و نمایاننده می باشد. از نظر کاربردی شاخص کمیتی است که نماینده چند متغیر همگن است (امیری، ۱۳۸۷). معیار مقایسه مقدار شاخص در

هر زمان نسبت به تاریخ مبدأ، عدد مبنا است. عدد مبنا به طور معمول ۱۰۰ در نظر گرفته می‌شود و در این صورت از تقسیم میزان شاخص در هر زمان بر عدد مزبور می‌توان رشد پارامتر مورد نظر در میان گروه متغیرهای مورد بررسی را به دست آورد. به عنوان مثال در صورتی که شاخص کل قیمت بورس تهران در تاریخ ۱۳۹۴/۱۲/۱۲ برابر ۷۷۶۹۷۸۰ باشد، مشخص است که سطح عمومی قیمت شرکت های بورس نسبت به فروردین ۱۳۶۹ که تاریخ مبدأ شاخص کل قیمت است، ۷۷۶۹۷/۸۰ برابر شده است. بدین ترتیب مشخص است که اجزا بنیادین هر شاخص را عوامل زیر تشکیل می دهند:

۱. گروه متغیرهای مورد بررسی مزبور؛

۲. پارامتر مورد نظر (به طور معمول قیمت)؛

۳. تاریخ مبدأ؛

۴. عدد مبنا.

عمده تفکیک شاخص ها بر حسب دامنه شمول شان و پارامتر مورد بررسی آنها صورت می گیرد. از نظر دامنه شمول همانگونه که متغیرهای مورد بررسی یکی از اجزای بنیادین هر شاخص است، این متغیرها به عنوان دامنه شمول آن شاخص محسوب می‌شود و قابل تفکیک هستند. بدین ترتیب شاخص های بخشی، شاخص شرکت های منتخب، شاخص های مرکب و دیگر شاخص ها هر یک روند عمومی قیمت سهام گروه خاصی از شرکت ها را نشان می دهند. پارامتر مورد بررسی یکی دیگر از اجزای بنیادین شاخص است که به طور معمول دو پارامتر قیمت و درآمد کل را مورد بررسی قرار می دهند. علاوه بر شاخص قیمت و شاخص بازده کل بعضی از بورس ها به طراحی و محاسبه شاخص بازده نقدی نیز می پردازند که پارامتر مورد بررسی آن، میزان بازده نقدی پرداختی شرکت های مشمول شاخص است. شاخص های بورس اوراق

بهادار تهران شاخص قیمت که از فروردین سال ۱۳۶۹ اقدام به محاسبه و انتشار آن به نام تپیکس^۱ شد. نحوه محاسبه تپیکس به صورت زیر است.

$$\text{TEPIX} = \frac{\sum_{i=1}^n Pit.Qit}{\sum_{i=1}^n Pio.Qio} \quad \text{رابطه (۲)-}$$

(۱)

P_{it} : قیمت سهام شرکت i ام در زمان t است.

Q_{it} : تعداد سهام شرکت i ام در زمان t است.

P_{i0} : قیمت سهام شرکت i ام در اولین دوره پذیرش است.

Q_{i0} : تعداد سهام شرکت i ام در اولین دوره پذیرش است.

از دیگر شاخص ها، شاخص صنایع که بر طبق ISIC به تفکیک ۲۸ صنعت و یا به طورت کلی تر در دو حوزه صنعت و مالی، شاخص پنجاه شرکت فعال تر، شاخص قیمت و بازده نقدی و شاخص بازده نقدی است. به اعتبار چشم انداز بیست ساله ی کشور، اصل چهل و چهار و جهت گیری کلی کشور مبنی بر واگذاری امور دولتی؛ راه را بر این هموار می سازد تا نیروهای بالقوه غیر دولتی وارد عمل شده و به تصدی امر بپردازند. که یکی از شاخصه های این موضوع واگذاری سهام شرکت های دولتی به عموم مردم از طریق بازار سرمایه می باشد. بنابراین آشنایی با عوامل مالی مؤثر بر شاخص قیمت سهام ضرورت پیدا می کند، تا بورس اوراق بهادار به عنوان مکانی سودآور برای مشارکت عمومی در توسعه اقتصادی ملی تبدیل گردد.

^۱ TEPIX

۲-۴- بازار کارا و امکان پیش‌بینی قیمت‌ها:

در اغلب مواقع از یوجین فرانسیس فاما، اقتصاددان امریکایی و یکی از برندگان نوبل اقتصاد سال ۲۰۱۳ میلادی، به عنوان پدر فرضیه بازار کارا یاد می‌شود که مطالعه در این زمینه را از دوره دکترایش آغاز کرده بود. او در ماه می سال ۱۹۷۰ میلادی مقاله ای علمی درباره نظریه مالی با عنوان "بازارهای سرمایه کارا: گذری بر نظریه و کارهای تجربی" منتشر کرد و برای نخستین بار فرضیه بازار کارا را بیان نمود. موضوع کارایی بازار یکی از مباحث مهم در زمینه سرمایه گذاری محسوب می‌شود. اکثر اقتصاددانان در حوزه امور مالی اعتقاد دارند که سرمایه را باید به نحو بهینه تخصیص داد. به عبارت دیگر هدف منطقی سیاست دولت‌ها، ترغیب و تشویق استقرار بازار کارایی تخصیصی است. ک در آن شرکت‌هایی که دارای محتمل‌ترین فرصت‌های سرمایه گذاری هستند، به وجوه مورد نیاز دسترسی داشته باشند. مفهومی که از کارایی در اینجا مدنظر قرار می‌گیرد، اشاره به این مسئله دارد که تا چه میزان بازار تعیین قیمت اوراق بهادار موفق عمل کرده است. بازاری از نظر تخصیص سرمایه کارا محسوب می‌شود که هم کارای درونی باشد و هم کارای بیرونی.

در یک بازار کارای بیرونی (که کارایی اطلاعاتی نیز نامیده می‌شود)، اطلاعات به سرعت و به طور وسیع منتشر شده و بدین وسیله امکان تعدیل سریع قیمت هر ورقه بهادار در مقابل اطلاعات جدید فراهم می‌شود به طوری که قیمت آن، ارزش سرمایه گذاری را منعکس می‌کند. بازار کارای درونی (که کارایی عملیاتی نیز نامیده می‌شود)، بازاری است که در آن کارگزاران و معامله‌گران به صورت منصفانه با یکدیگر رقابت می‌کنند به طوری که هزینه مبادلات کم و سرعت انجام آن زیاد است. به بیان بهتر کارایی درونی بر این موضوع تأکید دارد که کارگزاران و معامله‌گران با هزینه‌ای کم و به سرعت معامله را برای مشتری اجرا کنند. اصطلاح کارایی بازار معمولاً به کارایی بیرونی بازار دلالت دارد.

بازار کارا ویژگی های زیر دارد:

تعداد زیادی سرمایه گذار منطقی که به دنبال حداکثر کردن سود خود می باشند، به طور فعال از طریق تجزیه و تحلیل، ارزشیابی و معامله سهام در بازار مشارکت دارند. این سرمایه گذاران به تنهایی نمی توانند بر قیمت اوراق بهادار تأثیر بگذارند.

۱- کلیه سرمایه گذاران بدون متحمل شدن هزینه ایی به اطلاعات جاری موجود در مورد آینده دسترسی دارند.

۲- کلیه سرمایه گذاران قیمت های بازار را به دقت زیر نظر دارند و به طور مناسبی اوراق بهادار نگهداری شده را تعدیل می کنند.

۳- اطلاعات به صورت تصادفی به وجود می آید و مستقل از اعلان سایر اطلاعات می باشد.

۴- سرمایه گذاران در قبال ارائه اطلاعات جدید سریعاً عکس العمل نشان می دهند که این خود باعث تغییر قیمت سهام می شود.

در حقیقت می توان بیان کرد که بازار کارا بازاری است که در آن قیمت اوراق بهادار در کلیه زمان ها با ارزش سرمایه گذاری آن برابر است. مفهوم بازار کارا بر این فرض استوار است که سرمایه گذاران در تصمیمات خرید و فروش خود، تمامی اطلاعات مربوط را در قیمت سهام لحاظ خواهند کرد. بنابراین قیمت فعلی سهام شامل تمامی اطلاعات شناخته شده اعم از اطلاعات گذشته (مانند سود مربوط به فصل یا سال گذشته) و اطلاعات فعلی است.

۲-۵-اشکال مختلف کارایی بازار

۲-۵-۱-شکل ضعیف کارایی:

اگر برای به دست آوردن سود های غیر نرمال (بجز آنهایی که بطور شانسی به دست می آیند) با استفاده از قیمت های گذشته برای تصمیم گیری در زمان خرید و فروش اوراق بهادار امکان پذیر نباشد، عنوان می شود که بازار دارای شکل ضعیف کارایی است. به عبارتی می توان گفت در این نوع از کارایی بازار تغییرات گذشته قیمت باید رابطه ایی با تغییرات آنی قیمت نداشته باشد. به عبارت دیگر، تغییرات قیمت سهام در طول زمان باید مستقل و یا تقریباً این گونه باشد. مفهوم شکل ضعیف بازار کارا این است که اطلاعات قیمتی تاریخ گذشته در ارزیابی آتی قیمت تأثیر ندارد.

۲-۵-۲-شکل نیمه قوی کارایی

اگر کسب سودهای غیرنرمال (بجز آنهایی که بطور شانسی به دست می آیند) از طریق استفاده در دسترس عموم برای تصمیم گیری در مورد خرید و فروش اوراق بهادار امکان پذیر نباشد. عنوان می شود که بازار دارای شکل نیمه قوی کارایی است. این سطح که سطح خیلی متداول کارایی بازار می باشد نه تنها شامل اطلاعات موجود قیمتی است، بلکه شامل تمام اطلاعات شناخته شده و در دسترس مانند اطلاعات مربوط به درآمد، سود تقسیمی، اعلان تجزیه سهام، پیشرفت های جدید در مورد محصولات، مشکلات تأمین مالی و تغییرات مربوط به حسابداری نیز می شود. بازار کارای نیمه قوی در برگیرنده شکل ضعیف و بازار کارا نیز هست و شامل سرمایه گذارانی است که از اطلاعات منتشره جدید استفاده نمی کنند و این سرمایه گذاران انتظار دارند با توجه به متوسط ریسک، بازده بیشتری بدست آورند.

۲-۵-۳- شکل قوی کارایی:

اگر کسب سود های غیر نرمال (به جز آنهایی که بطور شانس به دست می آیند) از طریق استفاده از هر گونه اطلاعات برای تصمیم گیری در مورد خرید و فروش اوراق بهادار به هیچ وجه امکان پذیر نباشد، گفته می شود که بازار دارای شکل قوی کارایی است. شکل قوی بازار کارا در بر گیرنده شکل ضعیف و نیمه قوی نیز هست و نشان دهنده بالاترین سطح کارایی بازار است.

۲-۶- امکان پیش بینی قیمت ها در بازار:

تغییرات قیمت اوراق بهادار در بازار کارا به صورت تصادفی است. در بازار کارا تمام سرمایه گذاران فوراً با ورود اطلاعات جدید واکنش نشان می دهند و قیمت سهام با این واکنش تغییر خواهد نمود. در این بازار اطلاعات جدید خبری شگفت انگیز است (اطلاعاتی که شگفت انگیز نیست قابل پیش بینی بوده، قبلاً پیش بینی شده است و اکنون کهنه تلقی می شود). اخبار شگفت انگیز می توانند خوب یا بد باشند، در بازار کارا اخبار خوب منجر به افزایش قیمت و اخبار بد منجر به سقوط قیمت خواهد شد. با اعلام افزایش در بازدهی سهام (برای مثال، افزایش در سود نقدی سهام)، قیمت سهام آن افزایش می یابد. در بازار کارا هرگونه افزایش یا کاهش در میزان بازدهی غیر قابل پیش بینی است. بنابراین تغییرات قیمت در بازار کارا کاملاً تصادفی است. تصادفی بودن تغییرات قیمت به معنای غیر منطقی بودن قیمت فعلی نیست. به دلیل ورود تصادفی اطلاعات، تغییرات قیمت به عنوان برآیند ورود اطلاعات تصادفی (که گاهی مثبت و گاهی منفی) است، تلقی می شود. به هر حال تغییرات در قیمت، نتیجه ارزیابی مجدد سرمایه گذاران از وضعیت آتی سهام و اخذ تصمیمات جدید در مورد خرید یا فروش سهام است. از این رو، تغییرات در قیمت سهام اگر چه تصادفی است، اما منطقی است. مدیریت مالی جدید با بررسی تئوری بازار کارا، وجود ناهنجاری هایی را در بازار کارا مورد تأکید قرار داده است. دو دسته ناهنجاری های غیر تقویمی^۱ و تقویمی در تئوری

¹ Non Calender Anomalies

بازار کارا بیان شده است. ناهنجاری های غیر تقویمی بی قاعدگی ها و بی نظمی هایی هستند که فرضیه بازار کارا را زیر سؤال برده و نمی توان آن ها را در قالب بی نظمی های فصلی طبقه بندی کرد. در این بی قاعدگی ها عامل زمان بر هم زننده ی بازار کارا نیست. ناهنجاری های غیر تقویمی به پنج دسته اثر تجزیه سهام، اثر عرضه ی اولیه اوراق، اثر شاخص، اثر معاملات محرمانه و اثر تأخیر در ارائه گزارش سودآوری تقسیم می شوند.

اثر تجزیه سهام^۱: به طور منطقی تجزیه سهام یک شرکت نباید اثری بر ارزش شرکت فوق بگذارد، لیکن شواهد و مطالعات فراوانی حاکی از آن است که تجزیه سهام یک بنگاه سبب افزایش قیمت سهام هم قبل و هم بعد انتشار اعلامیه ی تجزیه ی سهام می شود. این اثر را " فاما، ادسای و جاین"^۲ مورد مطالعه قرار داده اند.

اثر عرضه اولیه اوراق^۳: ظریف فرد و مهرجو در سال ۱۳۸۳ با تحقیقی تحت عنوان " قیمت گذاری سهام شرکت ها در اولین عرضه سهام آن ها به بورس " بازده کوتاه مدت سهام ۹۱ شرکت جدید ورود در طول سال های ۱۳۷۸ تا ۱۳۸۲ را بررسی کردند. نتایج این تحقیق نشان داد بازده کوتاه مدت این شرکت ها بیشتر از بازده بازار بوده است.

اثر شاخص^۴: جنگ هاون^۵ در سال ۲۰۰۵ با بررسی شاخص S & P 500 و نحوه تغییر قیمت سهام شرکت هایی که در این شاخص قرار می گیرند و یا خارج می شوند، اظهار کرد که قیمت سهام شرکت هایی که به شاخصی همچون S & P 500 اضافه شوند به سرعت افزایش می یابد و بازده غیر معمول کسب می کنند.

¹ Stock Split Effect

² Fama, Dessai & Jain

³ Initial Public Offering Effect

⁴ Index Effect

⁵ Jong Hawan

اثر معاملات محرمانه^۱: افرادی نظیر مدیران شرکت، سهامداران عمده و برخی از افراد حرفه ای بازار، به دلیل عدم کارایی اطلاعاتی بازار، اطلاعات بروزتری نسبت به شرکت دارند. مطالعات زیادی همچون فیزیکی (۱۹۷۶)، سیپرن (۱۹۸۶)، لین هار (۱۹۹۰) و لاکونیشوک (۲۰۰۱) نشان دادند که محرمین اطلاعات شرکت اغلب می‌توانند تغییرات را پیش‌بینی کنند و بازده های غیر عادی نسبت به بازده بازار کسب کنند. اثر تأخیر در ارائه ی گزارش سودآوری: برخی محققان معتقدند شرکت هایی که گزارش سودهای خود را به سایر شرکت ها با تأخیر ارائه می دهند، در دوره زمانی قبل از ارائه ی گزارش با بازده های منفی مواجه می‌شوند. هر چند از نظر سودآوری تغییرات عمده ای در شرکت صورت نپذیرفته باشد.

ناهنجاری های تقویمی: در مورد الگو و بی قاعدگی های فصلی در بازار های مالی از جمله بازار سهام و اوراق قرضه، شواهد و مدرک زیادی از نیم قرن گذشته تا کنون در دست است. مطابق فرضیه ی بازار کارا، قیمت سهام در یک بازار کارا همواره به صورت تصادفی تغییر می یابد که دلیل آن نیز پاسخ مثبت سهام به اطلاعاتی است که به گونه ای تصادفی در طول زمان منتشر و عرضه می‌شود. حال اگر زمان، عامل تغییر دهنده قیمت سهام باشند، به گونه ای که طول دوره ی خاص، علاوه بر اطلاعات عرضه شده (به صورت تصادفی)، زمان نیز بر تغییر قیمت سهام مؤثر باشد و ماهیت تصادفی بودن بازار کارا را تغییر می دهد. به این گونه اثرات، اثرات تقویمی یا بی قاعدگی های تقویمی گفته می‌شود. این ناهنجاری ها به پنج دسته اثر روزهای آخر هفته، اثر ماه های خاص سال در تقویم اسلامی (ماه های رمضان، محرم و صفر)، اثر روزهای تعطیل با اثر قبل از روزهای تعطیل، اثر آخر ماه و بین ماه و اثر پایان سال (اثر ژانویه، اثر دسامبر) .

اثر روزهای آخر هفته: اثر روزهای هفته به وجود الگوهایی در بازده سهام در گذشته اشاره دارد که این الگو به روزهای ویژه هفته مربوط می‌شود. چنین روابطی به طور عمده در امریکا تصدیق شده است. به نحوی

^۱ Insider Transaction Effect

که آخرین روزهای معاملاتی هفته با بازدهی غیر عادی مثبت همراه است. در حالی که دوشنبه یعنی اولین روز معاملاتی هفت با بازدهی کمتر از روز های دیگر و حتی بازدهی منفی مشخص می شود. وجود چنین اثری خود می تواند شواهدی بر این ادعا باشد که بازدهی در روزهای مختلف هفته مستقل نبوده و در مقابل تئوری گشت تصادفی قرار می گیرد. بدری و صادقی به این نتیجه رسیدند که اثر چهارشنبه (اثر آخر هفته) و نیز یکشنبه در اکثر تخمین ها معنادار بوده است (ابدورضاک، ۲۰۱۴).

اثر ماه های خاص سال در تقویم اسلامی (ماه های رمضان، محرم و صفر): اثر ماه های خاص سال در تقویم به عنوان نمونه هایی از بی قاعدگی های تقویمی بازار مورد توجه بسیاری از محققان قرار گرفته است که در آن شواهدی دال بر وجود بازده های واقعی و غیر عادی یافت می شود. در این ماه ها هم زمان با کاهش روند رو به رشد فعالیت های اقتصادی و عملیات سفته بازی مسلمانان، تعداد معاملات در بازار های مالی کاهش می یابد. گرکز و اکرمی به بررسی ماه رمضان بر بازده (واقعی و غیر عادی) کردند. نتایج آنان حاکی از آن بوده که رابطه معناداری بین ماه رمضان و بازده واقعی وجود ندارد ولی با بازده غیر عادی سهام رابطه معناداری وجود دارد. معین الدین و عظیمی به بررسی تأثیر ماه محرم و صفر بر بازده سهام شرکت های پذیرفته شده در بورس پرداختند. نتایج حاکی از آن بود که بین ماه محرم و صفر و بازده سهام رابطه معنا داری وجود دارد (مهرآذین و اکرمی، ۲۰۱۲).

اثر روزهای تعطیل با اثر روزهای غیر تعطیل، مطالعات زادی در کشور های مختلف انجام شده که در آن شواهدی دال بر وجود بازده های غیر عادی بالا در روزهای قبل از تعطیلات رسمی یافت می شود. لیانو (۱۹۹۳) شواهدی از این امر را در بازار های خارج از بورس نشان داد. بیشتر این مطالعات در بازار های سهام امریکا انجام شده و اکثر آن ها چنین اثری، یعنی بازده های غیر عادی طی چند روز قبل از تعطیلات رسمی را تأیید کرده اند.

اثر آخر ماه و بین ماه: مطالعات اولیه مربوط به اثر بین ماه توسط آریل (۱۹۸۷) صورت پذیرفت. وی دریافت که به طور متوسط سهام شرکت ها در نیمه اول ماه (که از چند روز قبل از آخر ماه قبلی شروع می شود) بازده مثبت غیر عادی دارند. همچنین بازده عادی نیمه اول ماه به طور معناداری بیشتر از نیمه دوم ماه است. همچنین آرساد و کوثر (۱۹۹۷) نیز چنین اثری را در بازده های ۶۰ روزه در بازار سهام لندن نشان دادند.

اثر پایان سال (اثر ژانویه یا دسامبر): اثر ماهی سال مخصوصاً اثر آخر سال مورد توجه بسیاری از محققان قرار گرفته است و اصلی ترین حوزه بررسی در مطالعات کاربردی بی قاعدگی های بازار را به خود اختصاص داده است. طی سال های گذشته مطالعات زیادی بیان گر آن بوده اند که بازدهی سهام، به خصوص سهام شرکت های کوچک در نخستین ماه سال میلادی نسبت به سایر ماه ها به طور معناداری بیشتر است. این پدیده ای که در بسیاری از بازار های جهانی مورد بررسی قرار گرفته است، به اثر ژانویه (فروردین) معروف است. همچنین آن چه در آخرین ماه سال در بازار های مالی بسیاری از کشورها دیده شده و مورد بررسی قرار گرفته، وجود پدیده ای در جهت عکس اثر ژانویه است که به اثر دسامبر (اسفند) معروف است. این اثر نیز بیان می دارد میانگین بازدهی سهام در آخرین ماه سال نسبت به سایر ماه ها کمتر است.

وجود اثرات تقویمی بر بازدهی بازار سهام، اقتصاددانان مالی را متحیر کرده است. براساس فرضیه بازار کارا، الگوی فصلی نباید دارای اثرات معنادار باشد. اثرات تقویمی باعث بازدهی هایی می شوند که متناسب با ریسک نیستند. به عبارت دیگر، وجود اثرات تقویمی شکل ضعیف فرضیه بازار کارا را خنثی می کند.

در ایران پژوهش های بسیاری در زمینه ارزیابی میزان کارایی بازار بورس اوراق بهادار تهران انجام گرفته که در ادامه به برخی از آنها اشاره شده است. با توجه به تحقیقات زهرا نصراللهی در سال ۱۳۷۱ بازار بورس اوراق بهادار تهران در سطح ضعیف کارا می باشد. پژوهش اسماعیل فدائی نژاد در سال ۱۳۷۳ دلالت بر

عدم کارایی بورس اوراق بهادار تهران در سطح ضعیف دارد. حسنعلی سینایی در سال ۱۳۷۳ در سطح نیمه قوی کارایی بوررس اوراق بهادار تهران را بررسی کرده است و نتیجه تحقیق رد فرضیه کارایی بورس در سطح نیمه قوی بود و محمد نمازی و زکویه شوشتری در سال ۱۳۷۵ در مورد کارایی بازار در سطح ضعیف تحقیق کرده اند که نتایج این آزمون نیز عدم کارایی در سطح ضعیف بورس اوراق بهادار تهران را نشان داد. نتایج حاصل از تحقیق حسنعلی سینایی در سال ۱۳۸۴ وجود روندی آشوبناکی و غیر تصادفی در قیمت های سهام، به دلیل همبستگی بین قیمت های متوالی در سری زمانی نشان داد. علاوه بر آن طبق نتایج حاصل از پژوهش جواد عظیمی و همکاران در سال ۱۳۹۱، به دلیل وجود اثرات غیر عادی تقویمی از جمله فروردین، اثر ماه محرم و صفر، اثر رمضان و سایر اثرات ماهانه، الگوی خاصی در رفتار بازار وجود دارد که شناسایی و بهره گیری از روند این الگوها، سرمایه گذاران حرفه ایی را در جهت کسب بازده غیر عادی یاری می رساند، لذا می توان گفت بورس اوراق بهادار تهران حتی در سطح ضعیف یا نیمه قوی کارا نبوده و رفتار قیمتی سهام تابع الگوی تصادفی صرف نمی باشد و فرضیه گشت تصادفی قیمت ها رد شد و قیمت ها و بازده سهام در این بازار قابل پیش بینی می باشند.

۲-۷- انواع روش های پیش بینی

موضوع شناخت و بررسی رفتار قیمت اوراق بهادار و ارزشیابی دارایی های مالی، از آغاز شکل گیری بازارهای سرمایه همواره مورد توجه محافل علمی و سرمایه گذاران بوده است و روش های سنتی و مدرن بسیاری برای نیل به این هدف ارائه شده است. از جمله مکاتب مالی در خصوص پیش بینی، رویکرد تکنیکی، بنیادی، رویکرد مبتنی بر نظریه های مدرن مالی و رویکرد بی نظمی و پویایی غیرخطی است (کیم و لی، ۲۰۱۱).

۲-۷-۱-روش های کلاسیک پیش‌بینی

قبل از ارائه رایانه ها و استفاده از آن ها برای پیش‌بینی در حوزه های مالی، کار پیش‌بینی با روش های دیگری انجام می شده که در ادامه به برخی از آنان اشاره شده است.

۲-۷-۲-رویکرد تجزیه و تحلیل بنیادین:

این رویکرد و مدل‌های مورد استفاده در آن از دهه ۱۹۳۰ مطرح بوده، اما عمدتاً بعد از جنگ جهانی دوم در قالب نظری فراگیر، مورد توجه قرار گرفته اند. در این گونه مدل ها اساساً به ارزش ذاتی سهم توجه می‌شود. این دسته از روش ها به دلیل تعیین ارزش ذاتی سهم به صورت علمی و با تکیه بر ابزارهای مختلف علمی از قبیل اقتصاد، آمار، اطلاعات مالی و... مورد تأیید دانشگاهیان می باشد. مدل های مبتنی بر تحلیل بنیادی برای تعیین ارزش ذاتی سهم به صورت های مالی، سوابق تقسیم سود، سیاست های مدیریت، رشد فروش، توان مؤسسه در افزایش سودآوری و بسیاری عوامل دیگر توجه نموده، سپس ارزش ذاتی به دست آمده را با قیمت جاری سهم مقایسه می کنند. براین اساس در مورد خرید، فروش و یا نگهداری آن تصمیم گیری می نمایند. در میان روش های بنیادین، آن دسته نظریات که ارزش اتی سهام در توان ایجاد درآمد شرکت جستجو می نمایند را مورد توجه قرار داده است زیرا با اصول فرضیه بازار کارا نیز تطابق بیشتری دارد.

این رویکرد برای تعیین ارزش ذاتی سهام از روشی تحت عنوان تجزیه و تحلیل اساسی، پایه ای یا بنیادین استفاده می‌شود که در واقع همان بکار گیری تمامی اطلاعات موجود و در دسترس است. در این روش وضعیت یک شرکت را با توجه به تمام ابعاد اصلی خرد و کلان، مورد توجه و ارزیابی قرار می دهند. به بیان دیگر، ابتدا اوضاع اقتصادی کشور را مد نظر قرار داده، وضعیت شاخص های با اهمیتی مثل GDP، نرخ تورم، نرخ ارز و به طور خلاصه سیاست های پولی و مالی دولت را بررسی می نمایند. سپس برای تعیین وضعیت صنعت مورد نظر، به نوع تولید، مصرف کنندگان و تقاضای عمومی، وجود کالاهای

جانشین، حمایت های دولت، سیاست های قیمت گذاری و ثبات فروش توجه می کنند. در مرحله سوم به تجزیه و تحلیل شرکت پرداخته و پارامترهایی را از قبیل مدیریت کیفیت، کارایی عملیاتی، قابلیت سودآوری و بررسی ساختار مالی آن، مورد نظر قرار می دهند. سپس سهام شرکت ها را به طور مجزا بر مبنای ریسک و بازده بررسی و انتخاب می شوند و نهایتاً سهام باقی مانده با پرتفویی از دارایی ها ترکیب می شوند (اندرسون و همکاران، ۲۰۱۴).

۲-۷-۳- رویکرد تجزیه و تحلیل تکنیکال

از اوایل قرن بیستم که به تدریج رفتار قیمت سهام و ارزش آن به شکلی علمی تر مورد توجه قرار گرفت، برخی از دست اندرکاران و شرکت های سرمایه گذاری از طریق تعقیب قیمت و روندهای خاص، الگوی تغییرات قیمت را به دست آورده و نتایج کارهای خود را مبنای تصمیمات سرمایه گذاری قرار می دادند. ترسیم رفتار قیمت، بررسی و تهیه نمودارها و مطالعات نوسانات و شناخت حساسیت های رفتار قیمت و پیش بینی آینده، هدف اصلی این گروه از صاحب نظران می باشد. هنوز نیز این تفکر مورد قبول بسیاری از سرمایه گذاران و مؤسسات می باشد. این گروه را تحلیلگران تکنیکی یا چارتیست می خوانند، زیرا از منحنی ها و نمودارها بسیار استفاده می نمایند. این تحلیل گران معتقدند که عوامل مؤثر بر عرضه و تقاضا بی شمارند و هیچ گاه نمی توان آن ها را به دقت و بدرستی شناسایی نمود، لذا بهترین شیوه کار را مطالعه حرکات گذشته و به دست آوردن الگوی تغییرات آینده می دانند. آنان عرضه و تقاضا را وابسته به عوامل بسیار زیادی دانسته و اعتقاد دارند که قیمت های گذشته، منعکس کننده آینده بوده و قیمت را تابع محض عرضه و تقاضا می دانند. آنان به دنبال تغییرات بلند مدت نیستند و می گویند باید از فرصت های کوتاه مدت حداکثر استفاده را نمود و سود آبی به دست آورد. این روش در حال حاضر علاقمندان بسیار زیادی در میان تحلیلگران بازار دارد ولی به علت ضعف استدلال و توجیهات علمی، در مجامع دانشگاهی

مورد قبول و حمایت نیست (ادوارد و همکاران ۲۰۰۷). شکل ۲ نمونه ای از نمودارهای مورد استفاده در تحلیل تکنیکال است.



شکل ۲-۱: نمونه ای تحلیل قیمت سهام با ابزارهای رویکرد تحلیل تکنیکال

۲-۷-۴- رویکرد مبتنی بر نظریه های مدرن مالی:

پس از جنگ جهانی دوم، مجموع نظریه های جدیدی به جامعه مالی عرضه شد که عنوان نظریه های مالی را به خود اختصاص داد. دیدگاه های ارائه شده در این مطالعات، با بررسی های گذشته، تفاوتی اساسی داشت و در آن زمان مورد استفاده وسیع دست اندرکاران بازارهای سرمایه قرار گرفت. عمر نظریه های مدرن در حدود نیم قرن است. ظرف این مدت این نظریه ها مبنای محاسبه ارزش دارایی های مالی و پیش بینی قیمت اوراق بهادار بوده و کوشیده اند وضعیت بازار سرمایه را توضیح دهند. سه حوزه به هم پیوسته فرضیه بازار کارا، نظریه پرتفوی و مدل قیمت گذاری دارایی های سرمایه ای^۱ که از حوزه های این رویکرد هستند، مورد تأیید محافل دانشگاهی هستند. این نظریه ها بر بازارهای مالی جهان تأثیر عمیق گذاشته و پایه فکری مدیران مالی و سرمایه گذاری بوده و در سال های گذشته مرکز ثقل رد یا قبول سرمایه گذاری برای بسیاری از پروژه ها بوده اند.

^۱ CAPM: Capital Asset Pricing Model

برای اولین بار پیتر برن اشتاین^۱ در کتاب انگاره های سرمایه پیچیده تر شدن فناوری و عدم امکان رسم نمودار و درک پیام نوسانات بازار را مطرح کرد و اظهار کرد که دیگر نمی توان مانند کولی ها از روی برگ چای، آینده قیمت ها را پیشگویی کرد. این دیدگاه و نتایج حاصل از برخی مطالعات موجب شد تا نظریه پردازان، به بررسی فرضیه هایی روی آورند که در آن قیمت سهام به طریق تصادفی حرکت می کرد. این مفهوم علمی، گشت تصادفی^۲ نامیده شد و بتدریج در نشریه ها و کتابهای معتبر جهان خود را باز کرد و بر رویکرد مبتنی بر نظریه های مدرن مالی تأثیر گذاشت (برنستین، ۲۰۱۳).

مدل هایی که در ادامه به آن ها اشاره خواهد شد، زیرمجموعه ایی از مدل های اقتصاد سنجی محسوب می شوند. ویژگی های پیش بینی اقتصاد سنجی استفاده از متغیرهای متعدد اقتصادی و اجتماعی است که به طور علی با متغیر وابسته ارتباط دارند. این مدل ها متشکل از یک یا چند معادله توصیف کننده رابطه بین متغیرهای مختلف هستند که سعی دارند تا رابطه میان متغیرهای وابسته و مستقل را با استفاده از یک یا چند معادله رگرسیون، توصیف نمایند. در حقیقت اقتصادسنجی روش هایی برای شناسایی و تخمین مدل هایی با چند مجهول را ایجاد می کند. این متدها به محقق اجازه می دهند که استنتاجی علی - معلولی در شرایطی غیر شرایط آزمایشی کنترل شده ارائه دهد. مدل های اقتصاد سنجی گستره ی وسیعی را به خود اختصاص داده اند که در نگاهی اجمالی می توان آن ها را با توجه به تعداد عواملی که با متغیرهای وابسته رابطه علی دارند، به دو دسته مدل های تک عاملی و مدل های چند عاملی تقسیم نمود. در ادامه به تفصیل به توضیح این مدل ها پرداخته شده است.

۲-۸- مدل قیمت گذاری دارایی های سرمایه ایی (CAPM)

این مدل که یکی از معروف ترین مدل های شناسایی عوامل تأثیر گذار بر بازده سهام و پیش بینی آن محسوب می شود، توسط شارپ (۱۹۶۴)، لینتر (۱۹۶۵) و موسین (۱۹۶۶) براساس چهارچوب مارکوویتز

^۱ Peter Bernstein

^۲ Random Walk

ارائه شده است. این مدل فرض می کند که سرمایه گذاران از منطق مارکویتز در تشکیل پرتفو استفاده می کنند. فرض دیگر آن وجود یک دارایی (دارایی بدون ریسک) است که بازدهی مشخصی دارد. این فرض در ارزش گذاری دارایی ها با نرخ کاهش مناسب که در هر مدل ارزش گذاری به کار می رود، بسیار راهگشا بوده و اهمیت فراوانی دارد. به بیان دیگر اگر بتوان نرخ بازدهی را از یک سرمایه گذاری به دست می آید، برآورد کرد، می توان این نرخ بازدهی برآورد شده را با نرخ بازدهی مورد انتظار آن که بر اساس مدل CAPM حاصل می شود را مقایسه کرده و تعیین کرد که آیا این دارایی کمتر، بیشتر و یا به درستی قیمت گذاری شده است (آنگ و چن، ۲۰۱۴).

پژوهش های مقطعی اولیه در مورد بازده سهام همچون نیکلسون (۱۹۶۰) به دلیل نمونه های کوچک استفاده شده برای انجمن آزمون های تجربی مورد توجه کافی قرار نگرفت، تا اینکه پایگاه داده های ^۱CRSP و ^۲Compustat به وجود آمد و پژوهشگران توانستند نمونه هایی به اندازه کافی بزرگ (و با کیفیت کافی) برای نتایج قابل اطمینان ایجاد کنند. در پی آن پس از چند سال ارائه مدل CAPM هیچ راه قابل اطمینانی برای آزمون پیش بینی های این مدل در مدل مورد متغیرهای همچون نسبت ارزش بازار به ارزش دفتری و نسبت قیمت به سود وجود نداشت.

در ادامه پژوهش های مهم انجام شده در رابطه با نقض مدل CAPM متغیرهای مورد بررسی، به تفکیک شرح داده شده اند. این موارد در ادبیات مالی با عنوان بی قاعدگی های بازار شناخته می شوند. در حقیقت بی قاعدگی های بازار^۳ نتایج پژوهش های تجربی هستند که تئوری مدون قیمت گذاری دارایی ها همخوانی ندارند. این بی قاعدگی ها نشان دهند نا کارآمدی بازار (فرصت های سود آوری) با کامل نبودن مدل قیمت گذاری دارایی مورد استفاده می باشد.

^۱ Center for Research in Security Prices

^۲ Computer Statistics

^۳ Market Anomalies

۱. نسبت سود به قیمت (E/P):

یکی از مطالعات اولیه که پیش‌بینی‌های CAPM را رد کرد، پژوهش باسو (۱۹۷۷) بود. باسو با استفاده از یک دوره زمانی از آوریل ۱۹۵۷ تا مارس ۱۹۷۱، نشان داد که سهامی که نسبت سود به قیمت بالایی داشتند (یا نسبت P/E پایینی داشتند) به صورت معناداری بازده‌هایی بیش از سهام با نسبت سود به قیمت پایین تحصیل کردند. پژوهش‌های بعدی که توسط جف، کیم و وسترفیلد (۱۹۸۹) انجام شد، علاوه بر تأیید این نتیجه نشان داد که بر خلاف آن چه که توسط برخی پژوهشگران ادعا می‌شود، اثر E/P نقض مستقیم مدل CAPM (که باور دارد تنها ریسک سیستماتیک باید مبنا قرار گیرد) می‌باشد.

۲. اندازه شرکت:

بنز (۱۹۸۱) نشان داد که سهام شرکت‌های با ارزش بازار پایین، بازدهی متوسط بالاتری از سهام شرکت‌های با ارزش بازار بالا دارند. سایر پژوهشگران نظیر باسو (۱۹۸۳) نشان دادند که اندازه متمایز از اثر E/P می‌باشد. شرکت‌های کوچک بازدهی بالاتری دارند.

۳. برگشت معکوس بازده در بلند مدت^۱:

دبوند و تالر (۱۹۸۵) بازندگان را به عنوان سهامی که سه تا پنج سال گذشته بازدهی‌های پایینی داشته‌اند و برندگان را به عنوان سهامی که بازدهی‌های بالایی را در خلال یک دوره‌ی مشابه داشته‌اند، تعریف کردند. نتیجه اصلی پژوهش این دو نشان داد که در خلال سه تا پنج سال آینده، بازندگان بازدهی متوسط بالاتری را نسبت به برندگان داشته‌اند. این حرکت بازدهی‌ها نسبت به معکوس شدن طی افق‌های بلند مدت (یعنی برنده شده بازندگان) هنوز هم یکی از

¹ Long – Term Return Reversal

موارد نقض CAPM می باشد. کپرا، لاکونیشوک و ریتر (۱۹۹۲) نشان دادند که تفاوت های ریسک سیستماتیک (β) در این زمینه برای توجیه مدل CAPM کافی نمی باشد.

۴. نسبت ارزش دفتری به ارزش بازار (B/M):

رزنبرگ، رید و لانستین (۱۹۸۵) نشان دادند که سهم های با نسبت ارزش دفتری به ارزش بازاری بالا به صورت معناداری بازدهی های بالاتری نسبت به سهام با نسبت ارزش دفتری به ارزش بازار پایین دارند. چان، هامائو و لاکونیشوک (۱۹۹۱) به نتایج مشابهی در بازار ژاپن دست یافتند.

۵. اهرم مالی^۱:

باندارای (۱۹۹۸) دریافت که شرکت های با اهرم مالی بالا (نسبت " بدهی به سرمایه " بالا) از شرکت های با اهرم مالی پایین در خلال دوره ی ۱۹۷۹-۱۹۸۴ به طور متوسط بازده بالاتری داشتند. به بیان بهتر می توان گفت که آزمون های تجربی اولیه CAPM پیش بینی محوری آن را مبتنی بر وجود رابطه ی خطی مثبت بین ریسک سیستماتیک (بتا) و بازده سهام، مورد تأیید قرار دادند، با این وجود نتایج مطالعات اخیر حکایت از این دارد که ضریب بتا به عنوان شاخص ریسک سیستماتیک، توان تشریح اختلاف میانگین بازده سهام را ندارد و غیر از بتا، سایر متغیرها در چارچوب مدل CAPM محلی از اعراب ندارند. از آن جمله می توان به اندازه شرکت، نسبت ارزش دفتری به ارزش بازار، نسبت سود به قیمت و اهرم مالی که در تبیین اختلاف بازده سهام نقش مؤثری ایفا می کنند، اشاره کرد. در واقع ریسک های بی شماری در ارتباط با شرکت وجود دارد. برخی از این ریسک ها شامل ریسک بازار، ریسک ورشکستگی و ریسک نقدینگی است، در حالی که در مدل CAPM تنها از یک عامل جهت تشریح مجموع ریسک ها استفاده می شود. از نظر

¹ Financial Leverage

منطقی، به نظر می رسد مدلی با عوامل گوناگون، ممکن است قدرت توصیف کنندگی و پیش‌بینی کنندگی بیشتری را فراهم کند. به عبارت دیگر، عوامل اضافه، اجازه می دهند ریسک‌هایی توزیع شوند که شرکت با آن مواجه است. عامل ریسک بازار صرفاً ابعاد گوناگون ریسک را تجزیه می کند و قادر به بیان، تشریح و توصیف اثر انواع ریسک بر بازده نمی باشد. طی سال های ۱۹۸۰ تا ۱۹۹۰ انحرافات و ناهنجاری های مدل CAPM آشکار گردید. به عقیده پژوهشگران، این ناهنجاریها به عنوان چالشی بر اعتبار CAPM در توانایی تشریح بازده مورد انتظار توسط عامل ریسک سیستماتیک (بتا) مطرح می‌شود (لولین^۱ و نگل^۲، ۲۰۰۶).

برخلاف همه این شواهد منفی، CAPM تا اوایل دهه ۱۹۹۰ هنوز هم به عنوان پیش فرض توسط اغلب تحلیل گران مالی به کار گرفته می شد.

۲-۹-مدل فاما و فرنچ:

یوجین فاما^۳ که رساله‌ی دکتری خود در سال ۱۹۶۵ به بررسی گشت تصادفی حرکت قیمت های سهام با استفاده از مفروضات معین اختصاص داده بود، در سال ۱۹۹۲، مقاله ای منتشر کرد که بیشتر پژوهش های تجربی قبلی را یک جا جمع آوری کرده بود. فاما و فرنچ (۱۹۹۲) متغیرهای اندازه، اهرم مالی، E/P، ارزش دفتری به ارزش بازار و بتا را در یک پژوهش مقطعی^۴ آزمون کردند. آن ها قدرت توضیح دهندگی اندازه، اهرم، E/P، نسبت ارزش دفتری به ارزش بازار و ریسک سیستماتیک را در رگرسیون مقطعی طی دوره ۱۹۶۳-۱۹۹۰ مقایسه کردند. نتایج نشان داد که اندازه شرکت و نسبت ارزش دفتری به ارزش بازار (ارزش شرکت) متغیرهایی هستند که قوی ترین رابطه را با بازده دارند و هنگامی که این دو متغیر در رگرسیون وارد می‌شوند، قدرت توضیح

¹ Lewellen

² Negel

³ Eyogin Fama

⁴ Cross Section Research

دهندگی سایر متغیرها ناپدید می‌گردد و بازده‌های متوسط مقطعی می‌توانند به صورت مفیدی توسط این دو متغیر توضیح داده شوند (گانت^۱، ۲۰۰۴).

نتایج فاما و فرنچ در سال ۱۹۹۲ این دیدگاه را که مدل تک‌عاملی CAPM روشی است که توسط آن اوراق بهادار به صورت واقعی قیمت‌گذاری می‌شوند، را خدشه‌دار ساخت. مدلی که بیش از هر مدل دیگر در دانشکده‌های مالی آموزش داده شده بود، به نظر میرسید که دیگر جواب‌گو نباشد. براساس این نتایج آن‌ها در سال ۱۹۹۳ مدل سه‌عاملی خود را ارائه کردند و نشان دادند که اندازه و نسبت ارزش دفتری به ارزش بازار دو عامل ریسک هستند که در CAPM نادیده گرفته شده‌اند. به عقیده آن‌ها، نسبت‌های قیمتی حاوی اطلاعات بسیار مشابهی در مورد بازده مورد انتظار هستند. آن‌ها به نتیجه مشابهی با استفاده از رویکرد رگرسیون سری زمانی در ارتباط با پرتفیلوهای سهامی دست یافتند که براساس قیمت مرتب شده بودند.

۲-۹-۱- عامل اندازه شرکت‌ها:

فاما و فرنچ به منظور تعیین اثر اندازه شرکت‌ها به ایجاد پرتفو‌هایی بر مبنای دو عامل اندازه و بتای سهام اقدام کردند. نتایج نشان داد که برای شرکت‌هایی که از اندازه در یک سطح قرار گرفته‌اند، رابطه معناداری بین عامل بتا و بازده سهام وجود ندارد و در واقع این عامل اندازه است که توانایی تبیین نوسانات بازده را در این شرکت‌ها دارا می‌باشد.

۲-۹-۲- عامل ارزش شرکت‌ها:

پژوهش‌های بسیاری از نسبت مبلغ دفتری به ارزش بازار حقوق صاحبان سهام برای ارزش‌گذاری این متغیر به عنوان یک عامل خطر برای توضیح نوسانات بازده سهام بهره‌برده‌اند. در این زمینه گریفمن و لمون ثابت کرده‌اند که اگر بازده یک سهم در قبال هر واحد تغییر در این نسبت

¹ Gaunt

حساسیت بیشتری از خود بروز داد، ریسک خرید آن سهام بیشتر است. از نظر فاما و فرنچ این نسبت در مقایسه با متغیرهای مشابه پارامتر ممتازتری برای توضیح و پیش‌بینی نوسانات بازده سهام به حساب می‌آید.

مدل سه عاملی فاما و فرنچ نیز در معرض انتقادهای فراوانی قرار گرفت که این انتقادها در چارچوب سه موضوع محوری داده کاوی، اشتباه انتخاب و روش برآورد ریسک سیستماتیک مطرح شد. مخالفت‌هایی که با پژوهش فاما و فرنچ شد و حمایت‌هایی که از آن صورت گرفت، یکی از جالب‌ترین شاخه‌های ادبیات علوم مالی می‌باشد. به ندرت یک مقوله دانشگاهی را می‌توان یافت که به این اندازه در دنیای واقعی کاربرد داشته باشد.

هدف اصلی و اولیه تجزیه و تحلیل اوراق بهادار نسبت به تغییرات آنها، یافتن تعداد عوامل مؤثر بر نرخ بازدهی اوراق می‌باشد. به بیان رسمی به چنین رابطه‌ای اصطلاحاً مدل عاملی نرخ‌هایی بازدهی اوراق بهادار گفته می‌شود.

این مدل‌ها به دو گروه مدل‌های تک‌عاملی و مدل‌های چندعاملی تقسیم می‌شوند. مفهوم اساسی در مدل تک‌عامل این است که تمامی اوراق بهادار از نوسانات عمومی بازار تأثیر می‌پذیرند، زیرا نیروهای اقتصادی مشابه در آینده اکثر شرکت‌ها را تحت تأثیر قرار می‌دهد. از مدل‌های تک‌عاملی می‌توان به مدل CAPM و مدل بازار اشاره کرد. و لیکن هدف اصلی در مدل‌های چندعاملی یافتن برخی از تأثیرات غیربازاری است که منجر به حرکت توأم سهام با یکدیگر می‌شود چرا که با مطالعه آزمون‌های تجربی مدل تک‌عاملی CAPM، مشاهده می‌شود که عوامل دیگری در توضیح بازده سهام نقش دارند. این موضوع منجر به شکل‌گیری مدل‌های چندعاملی شده است.

۲-۹-۳- سایر مدل های چند عاملی:

مدل های چند عاملی فرض می کنند که بازده اوراق بهادار به حرکات عوامل و یا شاخص های متعددی حساس است، در حالی که در مدل بازار فرض بر این است که اوراق بهادار صرفاً از شاخص بازار تبعیت می کند. به منظور برآورد صحیح تر بازده مورد انتظار، واریانس و همبستگی اوراق بهادار، مدل های چند عاملی توانایی بالاتری از مدل های بازار دارند. علت این امر آن است که بازده واقعی اوراق بهادار به عواملی بیش از صرف شاخص بازار حساس می باشد. به عبارت دیگر، در اقتصاد واقعی بیش از یک عامل تأثیر گذار بر بازده اوراق بهادار وجود دارد.

به عنوان یک فرآیند ایجاد بازده، مدل های چند عاملی سعی دارند تا نیروهای عمده اقتصادی که به طور سیستماتیک قیمت های کلیه اوراق بهادار را تغییر می دهند را کشف کنند. البته سرمایه گذاران نیز مدل عاملی را به طور ضمنی و یا صراحتاً در تصمیم گیری های خود به کار می گیرند. برای برآورد مدل های چند عاملی روش های مختلفی مورد استفاده قرار می گیرد که این روش ها را می توان در بدو امر به سه دسته طبقه بندی کرد: (۱) روش سری زمانی (۲) روش مقطعی (۳) روش تجزیه و تحلیل عامل^۱.

در روش سری زمانی فرض بر آن است که تحلیلگر عوامل تأثیرگذار بر بازده سهام را پیشاپیش می داند. شناخت عوامل مرتبط معمولاً از تجزیه و تحلیل های عوامل اقتصادی بر کل شرکت از جمله عوامل خرد و کلان اقتصادی، ساختار صنعت و تجزیه و تحلیل اساسی اوراق بهادار بدست می آیند.

در روش مقطعی ابتدا حساسیت اوراق بهادار به عوامل معین تخمین زده می شود و سپس در دوره زمانی خاصی ارزش عوامل مختلف بر مبنای بازده اوراق بهادار و حساسیت آن ها به این عوامل برآورد می شود.

^۱ Factor Analysis

این فرآیند طی چندین دوره تکرار شده به طوری که برآوردی از انحراف معیار عوامل هبستگی آن ها بدست می آید.

در روش تجزیه و تحلیل عوامل نه ارزش عوامل و نه حساسیت این عوامل مشخص است و از تکنیک آماری تجزیه و تحلیل عامل برای استخراج تعداد و حساسیت عوامل متعدد استفاده می شود. در هر صورت مدل برازش شده در مدل چند عاملی به شکل رابطه (۲-۷) می باشد:

$$R_{it} = a_i + b_{i1}F_{1t} + b_{i2}F_{2t} + \dots + b_{in}F_{nt} + e_{it} \quad \text{رابطه (۲-۲)}$$

در مدل فوق R_{it} بازده سهام i در زمان t و a_i مقدار ثابت و b_i برای هر سهم عددی ثابت و نشان دهنده حساسیت هر عامل، F عامل مشترک برای کلیه سهام و e_i جز تصادفی برای هر سهم می باشد. یکی از مدل های چند عاملی مدل کار هارت می باشد. این مدل شکلی از مدل سه عامله فاما و فرنچ است که کارهات در سال ۱۹۹۷ این مدل چهار متغیره را با اضافه کردن یک متغیر جدید تحت عنوان عامل شتاب ابداع کرد. مدل کارهات بعدها توسط سو و چن آزمون شد و نشان داد که این مدل چهار متغیره قدرت بیشتری در پیش بینی بازده نسبت به مدل سه متغیره فاما و فرنچ و مدل تک متغیره CAPM دارد. در میان مدل های چند عاملی، تکنیک های پیش بینی سری زمانی بیش از سایر روش ها توجه محققان را به خود جلب کرده است. تجزیه و تحلیل سری های زمانی با تئوری ارزش گاری دارایی ها در طول زمان گره خورده است. تفاوت مهمی که سری های زمانی مالی را از سایر سری زمانی جدا می سازد این است که تئری های مالی و بنیان سری های زمانی آن، با نوعی عدم اطمینان همراه است و به همین علت است که تئوری ها و روش های آماری نقش اساسی ای در تجزیه و تحلیل سری های زمانی مالی ایفا می کنند. سری های زمانی مالی انواع مختلف دارند و به شیوه های مختلفی قابل تقسیم بندی می باشند که مهمترین و پرکاربردترین آن ها عبارتند از:

الف- سری های زمانی خطی^۱:

این دسته از سری های زمانی که بنیان و ریشه سایر مدل ها را تشکیل می دهند، چهارچوب ساده ایی به منظور مطالعه ساختارهای پویا مهیا می سازند. مطالعه این مدل ها شامل وابستگی های ایستا و پویا، تابع همبستگی، مدل سازی و پیش بینی می شوند (میلز^۲ و مارکلوس^۳، ۲۰۱۰). برخی از این مدل ها به صورت زیر تقسیم بندی می شوند:

۱. مدل های خود توضیح رگرسیون ساده (AR)^۴

۲. مدل های میانگین متحرک ساده (MA)^۵

۳. مدل های تلفیقی خود توضیح میانگین متحرک ساده (ARMA)^۶

۴. مدل های فصلی^۷

۵. مدل های پویای تک ریشه ایی^۸

ب: سری های زمانی ناهموار شرطی^۹

این دسته از تکنیک ها به مدل سازی نوسانات و تغییر پذیری^{۱۰} بازده دارایی ها می پردازد. در این جا منظور از تغییر پذیری و نوسان، انحراف معیار شرطی بازده دارایی ها مورد بررسی است که در مدیریت

¹ Linear Time Series

² Mills

³ Markellos

⁴ Simple Autoregressive (AR) Models

⁵ Simple Moving – Average (MA) Models

⁶ Mixed Autoregressive Moving – Average (ARMA) Models

⁷ Seasonal Models

⁸ Unit – Root Nonstationary

⁹ Conditional Heteroscedastic Models

¹⁰ Volatility

ریسک دارایی ها کاربرد فراوان دارد و موجب بهبود تخمین پارامتر و افزایش دقت پیش‌بینی می‌شود (بانونس^۱ و لاورنت^۲، ۲۰۱۵). مهم ترین مدل معادلات تک متغیره تغییر پذیری عبارتند از

۱. مدل رگرسیون ناهموار خود توضیح شرطی^۳ (ARCH)

۲. مدل رگرسیون ناهموار خود توضیح شرطی تعمیم یافته^۴ (GARCH)

۳. مدل رگرسیون ناهموار خود توضیح شرطی تعمیم یافته نمایی^۵ (EGARCH)

۴. مدل رگرسیون ناهموار خود توضیح شرطی تعمیم یافته آستانه ایی^۶ (TGARCH)

۵. مدل رگرسیون ناهموار خود توضیح شرطی میانگین متحرک^۷ (CHARMA)

ج: سری های زمانی غیر خطی^۸

این دسته از مدل ها که با پیشرفت تکنولوژی روند رو به رشد داشته اند، توسط تست های پارامتری و غیر پارامتری آماری از سری های زمانی خطی قابل تمییز می باشند. مفاهیم بنیادی این دسته از سری های زمانی استفاده از مدل های شبیه سازی به منظور توصیف تکامل توزیع شرطی متغیر مربوطه و یا خصوصیات غیر خطی آن می باشد (زیوت^۹ و وانگ^{۱۰}، ۲۰۰۸). مهم ترین این مدل ها عبارتند از:

۱. مدل دوسویه^{۱۱}

¹ Banwens

² Laurent

³ Autoregressive Conditional Heteroscedastic (ARCH) Model

⁴ Generalized ARCH (GARCH) Model

⁵ Exponential GARCH (EGARCH) Model

⁶ Threshold GARCH (TGARCH) model

⁷ Conditional Heteroscedastic Autoregressive Moving Avrage (CHARMA)

⁸ None – Linear Time Series

⁹ Zivot

¹⁰ Wang

¹¹ Bilinear Model

۲. مدل رگرسیون خود توضیح آستانه ایی^۱ (TAR)

۳. مدل رگرسیون خود توضیح انتقال محور^۲ (STAR)

۴. مدل راه گزینی مارکو^۳

۵. مدل های غیر پارامتری^۴ مثل رگرسیون کرنل^۵

۲-۱۰- روش های مدرن پیش بینی

به دنبال تلاش های دانشمندان علوم ریاضی و سیستم های پویا روش های جدیدی را برای پیش بینی فعالیت های مالی ایجاد شده است. کاربرد مدل های غیر خطی و همچنین تکنیک های غیر پیشرفته در مدت زمان کمی توانسته است جایگاه خود را در علوم مختلف، به ویژه اقتصاد باز کند. متخصصین سیستم های غیر خطی، سعی در توضیح رفتار بازار های مالی و پیش بینی آن از طریق روش های پیشرفته غیر خطی کرده اند. این روش های عبارتند از نظریه آشوب و روش های هوش مصنوعی.

از آن جایی که پیش بینی های بازار های مالی، غیر خطی و غیر متریک هستند و از مدل های بی نظمی در طبیعت تبعیت می کنند، پیش بینی آنها از روش ها و متدهای متداول مانند سری زمانی و رگرسیون، به سختی می توان به پاسخ صحیح دست یافت. در این گونه مسائل مدل های شبیه سازی، داده کاوی، سیستم های خبره و روش شبکه های عصبی برای رفتارهای ناشناخته متغیرهای دارای نوسان بالا، ایده آل به نظر می رسد.

¹ Threshold Autoregressive

² Smooth Transition AR (STAR) Model

³ Markov Switching Model

⁴ Nonparametric Methods

⁵ Kernel Regression

۲-۱۰-۱- رویکرد بی‌نظمی و پویایی غیر خطی

از اواسط دهه ۱۹۷۰ و به ویژه از سال ۱۹۹۰ کوشش‌های گسترده‌ای و جدید در زمینه پیش‌بینی قیمت سهام، با استفاده از روش‌های جدید ریاضی، سری‌های زمانی طولانی و ابزار پیشرفته آغاز گردید، که منجر به ظهور دیدگاه بی‌نظمی و پویایی‌های غیر خطی شد. پس از آن مطالعات بسیاری بر روی اطلاعات قیمت و شاخص سهام در کشورهای دارای بازارهای مالی پیشرفته انجام شد تا وجود ساختاری معین در اطلاعات قیمت سهام نشان داده شود که از این طریق فرضیه گشت تصادفی را نقض کنند.

اندیشمندان در این چارچوب، در جستجوی روش‌های علمی جدیدتری هستند و معتقدند، نظریه بی‌نظمی و پویایی غیر خطی می‌تواند مقدمه ارائه نظریه‌های علمی تازه‌ای در مدیریت مالی باشد. رنگ باختن نظریه‌های مدرن مالی به همراه دو پدیده جهانی شدن بازارهای مالی و وجود قدرت‌های تکنولوژیک در سطح جهان، موجب شده است تا نظریه‌های مدرن مالی در بازارهای مختلف به سادگی مورد آزمون قرار گیرند و ناتوانی آنها در تبیین اقتصاد برای صاحبان اندیشه، روشن شود. بنابراین، تناقض‌های آشکار میان نظریه‌های موجود و دنیای واقعی، موجب شده است تا دانشگاهیان و اهل نظر، نسبت به اعتبار نظریه‌های مدرن مالی موجود، تردید کنند و نظریه بازار کارا و به ویژه مدل یک متغیره شارپ^۱ را به دیده تردید بنگرند.

گروهی از صاحب نظران و اندیشمندان قدم فراتر گذاشته و به پدیده آشفتگی^۲ توجه نمودند. آنها با استفاده از فنون جدید ریاضی، بازار راسیستم پیچیده و در حال تحول می‌دانند و معتقدند، اگر زمینه و جوهره بازار به درستی درک شود، هر قفلی در آن گشودنی است. این گروه مانند ریاضی دانان و فیزیک دانان، اعتقاد دارند چنانچه هر پدیده‌ای به درستی مشاهده و بررسی شود، وقایع به ظاهر تصادفی آن تبیین پذیر است و اگر در خور پیش‌بینی نباشد، حداقل قابل درک خواهد بود.

^۱ Sharp

^۲ Chaos Theory

طرفداران الگوی آشفتگی اذعان می کنند که رفتار قیمت ها به صورت غیر خطی است. حرکات قیمت سهام ناشی از عوامل بسیار زیادی است که در هر لحظه از زمان در آن ها تأثیر می گذارد. آنان در واقع معتقدند که به تعداد سرمایه گذاران، عامل تأثیرگذار بر قیمت سهام وجود دارد. در این راستا به دلیل وجود روابط غیر خطی پیچیده با متغیر های متعدد، توانایی و کنترل مدل های رگرسیون خطی به دلیل زمانبر بودن آنها در پردازش اطلاعات کاهش می یابد و لذا باید از یک سری روابطی مبتنی بر محاسبات نرم^۱ استفاده شود.

۲-۱۱- پیشینه ی پژوهش

اصلی ترین عامل که هر سرمایه گذار در تصمیمات خود مورد توجه خاص قرار می دهد، نرخ بازده است. یعنی سرمایه گذاران به دنبال پر بازده ترین فرصت ها برای سرمایه گذاری مازاد منابع خود در بازارهای سرمایه هستند. از طرفی در زمینه پیش بینی بازده سهام که یکی از موضوع های مورد علاقه سرمایه گذاران و محققان مالی است تا کنون تلاش های زیادی برای ارائه مدلی که بتواند بازده سهام را به نحو در خور اتکایی پیش بینی کند، انجام شده است (نمازی و کیا مهر، ۱۳۸۶).

پژوهش های اخیر توجه بیشتری به بازده سهام و پویایی دارند که به وسیله رفتار غیر خطی توصیف می شود چرا که بازارهای سهام رفتاری آشوب گون دارند و نه تصادفی. حرکت بازده و قیمت سهام در طول زمان از عوامل مختلفی تأثیر می پذیرد و اینگونه رفتار ایستا نبوده و حرکتی پویا را نشان می دهد، لذا مدل های خطی و استاتیک قادر به تبیین رفتار چنین سیستم هایی نیستند و به همین دلیل گرایش به استفاده از مدل های غیر خطی روز به روز افزایش یافته است.

^۱ Stof Computing

از میان مدل های غیر خطی، مدل های سری زمانی غیر خطی^۱ توجه بسیاری از محققان را به خود جلب کرده است و پژوهش های بیشماری در این زمینه انجام شده است. به منظور پیش بینی بازده سهام از مدل های خود توضیح آستانه ای، مدل خود توضیح انتقال هموار و ... به وفور استفاده شده است. دسته دیگر مدل های سری زمانی غیر خطی که پژوهش های بسیاری را به خود اختصاص داده اند، مدل شبکه عصبی مصنوعی است. این مدل به دلیل ویژگی های غیر خطی و ناپارامتری بودن و برخورداری از ساختاری داده محور، در تخمین مقادیر موفق تر عمل کرده است. تکنیک داده کاوی تکامل دستوری با توجه به مدت زمان کوتاهی که در تجزیه و تحلیل مسائل مالی کاربردی شده است، توجه بسیاری از محققان را به خود جلب کرده و عملکرد مناسبی با دقت بالا از خود نشان داده است. در ادامه به برخی از پژوهش ها در داخل و خارج از کشور اشاره شده است.

۲-۱۱-۱- پیشینه داخلی

در زمینه پیش بینی شاخص بورس، قیمت سهام و بازده سهام در ایران پژوهش های زیادی صورت گرفته است. در تحقیقات داخلی در این زمینه بیشتر از مدل های عصبی، عصبی - فازی، مدل های غیر خطی سری زمانی و الگوریتم ژنتیک استفاده گردیده است و از روشهای تکاملی بخصوص تکامل دستوری استفاده نشده است. بنابراین در این قسمت به بیان پیشینه پژوهش های مشابه که با سایر تکنیک ها انجام شده، پرداخته شده است که از این جمله می توان به پژوهش های زیر اشاره نمود.

علی رحمانی و فرشته سعیدی (۱۳۸۷) در پژوهش خود به منظور پیش بینی بازده سهام با استفاده از اطلاعات حسابداری و صورت های مالی از مدل لاجیت استفاده کردند و نتایج آن را با روش رگرسیون مقایسه نمودند. نتایج این پژوهش نشان داد که عملکرد دو مدل تفاوت معناداری ندارند.

^۱ Nonlinear Time Series Models

البراهیم عباسی و سحر باقری (۱۳۹۰) بازده سهام را با استفاده از مدل های غیرخطی خود توضیح آستانه ای (TAR) و خود توضیح انتقال محور لجستیک (LSTAR) پیش‌بینی نموده و با مدل خطی ARMA مقایسه کردند. نتایج پژوهش آنها دقت پیش‌بینی بیشتر مدل های غیر خطی نسبت به مدل ARMA را نشان داد.

نظیفی نایینی، فتاحی و صمدی (۱۳۹۱) با استفاده از مدل انتقالی گارچ و مارکف به پیش‌بینی نوسانات بازار بورس پرداختند. نتایج پژوهش آنها حاکی از آن است که برای پیش‌بینی نوسانات بازار سهام ایران، عملکرد مدل های SW-GARCH با توزیع خطای t و با درجه آزادی متغیر بین دو رژیم، بسیار بهتر از مدل های گارچ معمولی است. حتی در برازش و بررسی های داخل نمونه ایی نیز این مدل از مدل های انتقالی مارکف، رتبه اول را در زمینه قدرت به خود اختصاص داده اند.

زهرا پور زمانی و علی بشیری (۱۳۹۲) طی پژوهش خود در زمینه ارزیابی مدل کارهارت برای پیش‌بینی بازده مورد انتظار از مدل های رگرسیون چند متغیره و آزمون زوجی استفاده کردند.

رضا راعی و کاظم چاوشی (۱۳۸۲) در پژوهش خود به پیش‌بینی رفتار بازده سهام شرکت توسعه صنعتی بهشهر به وسیله مدل خطی عاملی (رگرسیون خطی چند متغیره) و شبکه های عصبی مصنوعی با معماری MLP و الگوریتم پس انتشار خطا پرداختند. متغیرهای مستقل انتخابی آنها که متغیرهای کلان اقتصادی هستند عبارتند از شاخص کل قیمت بورس تهران، نرخ ارز در بازار آزاد، قیمت نفت و قیمت طلا. نتایج این پژوهش ضمن نشان دادن توانایی هر دو مدل در پیش‌بینی رفتار بازده سهام، گویای برتری عملکرد شبکه عصبی مصنوعی بر مدل چند عاملی بود.

محسن مهر آرا و همکاران (۱۳۸۸) پژوهش الگوسازی و پیش‌بینی شاخص بورس اوراق بهادار تهران و تعیین متغیرهای مؤثر بر آن انجام دادند. در این مطالعه با الگوسازی و پیش‌بینی شاخص قیمت و بازده

نقدی بورس اوراق بهادار تهران مبتنی بر ساختار تلفیقی الگوریتم ژنتیک با رویکرد شبکه عصبی GMDH، سعی در شناخت متغیرهای مؤثر بر شاخص بورس اوراق بهادار شده است. با بکارگیری متغیرهای مختلف از جمله متغیرهای کلان اقتصادی در این تحقیق نشان از تأثیر قوی و معنادار شاخص قیمت زمین، هزینه مسکن، CPI، پایه پولی، کرایه مسکن اجاره ای و قیمت جهانی نفت خام بر شاخص قیمت و بازده نقدی بورس اوراق بهادار است. در مقابل، بازار ارز خارجی و طلا، ارتباط کمتری با بازار سهام داشته است.

حسنعلی سینایی و همکاران (۱۳۸۴) پژوهشی در زمینه پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه های عصبی مصنوعی و آرایه شواهدی مبنی بر رفتار آشوبناک شاخص قیمت در بورس اوراق بهادار انجام داده اند. در این مطالعه دو مجموعه از داده ها برای ورودی شبکه های عصبی به کار گرفته شده است. نتایج حاصل از پژوهش نشان از عملکرد مناسب در پیش‌بینی شاخص قیمت دارد و در رفتار شاخص قیمت حرکات آشوب ناک وجود دارد و همچنین با بکارگیری آزمون های مختلف شواهدی علیه فرضیه بازار کارا و گشت تصادفی وجود دارد.

فدایی نژاد (۱۳۷۳)، مطالعه ای را با استفاده از روشهای خود همبستگی و آزمون گردشها وبا بکارگیری قیمت هفتگی ۵۰ شرکت برای دوره زمانی ۱۳۶۸-۱۳۷۲ انجام داد و کارایی بازار بورس اوراق بهادار تهران را در سطح ضعیف ارزیابی کرد (فدایی نژاد، ۱۳۷۴).

اله یاری، اکبر (۱۳۸۷) در پژوهشی با استفاده از قیمت روزانه سهام ۹۵ شرکت بورسی در بازه زمانی ۱۳۸۴ تا ۱۳۸۷ به بررسی کارایی بازار بورس بوسیله روش های آماری تحلیل همبستگی و آزمون RUN پرداخته است. نتایج تحقیق بیانگر این است که تغییر قیمت سهام تصادفی نبوده و دارای روند قابل پیش‌بینی می باشد. بنابراین بورس تهران فاقد کارایی ضعیف می باشد (اله یاری & اکبر، ۱۳۸۷).

احمد عاملی و همکاران (۱۳۹۴) تحقیقی در زمینه پیش‌بینی قیمت سهام با استفاده از شبکه عصبی فازی مبتنی بر الگوریتم ژنتیک و مقایسه با شبکه عصبی فازی انجام داده‌اند. در این تحقیق که شامل دو بخش پیش‌پردازش و پیش‌بینی کننده است، در بخش اول سه فرآیند جای‌گذاری داده‌های غیر موجود، نرمالیزه کردن و انتخاب ویژگی به ترتیب صورت گرفته است. در بخش پیش‌بینی کننده با توجه به قابلیت پیش‌بینی هوشمند شبکه عصبی - فازی، از این بخش با دو ساختار ممدانی و سوگنو به عنوان پیش‌بینی کننده قیمت سهام استفاده شده است. با توجه به معیارهای مورد ارزیابی، نتایج به دست آمده برتری شبکه عصبی فازی ترکیبی را به شبکه عصبی فازی ساده نشان می‌دهد، اما به طور کلی پیش‌بینی کننده با ساختار سوگنو با الگوریتم ژنتیک دارای عملکرد بهتری نسبت به ساختار ممدانی دارد، چون تعداد پارامترهای آموزش ساختار سوگنو بیشتر است.

رضا تهرانی و همکاران (۱۳۹۱) پژوهش پیش‌بینی شاخص بورس اوراق بهادار با استفاده از مدل‌های شبکه عصبی مصنوعی شعاع پایه انجام داده‌اند. در این پژوهش عملکرد شبکه عصبی شعاع پایه و شبکه عصبی پرسپترون مقایسه شده‌اند. نوع آزمون عملکرد شبکه‌های عصبی براساس حداقل مربعات خطا در دو رویکرد درون نمونه‌ای برتری شبکه عصبی شعاع پایه و در رویکرد برون نمونه‌ای برتری شبکه عصبی پرسپترون را نمایش می‌دهد.

محمد رضا کریمی و همکاران (۱۳۹۰) مطالعه پیش‌بینی روند تغییرات شاخص کل قیمت سهام در بورس اوراق بهادار تهران با رویکرد فازی انجام داده‌اند. در این پژوهش از الگوریتم‌های بهینه و کارا در هوش مصنوعی استفاده شده است. این موضوع از آنجا مدنظر قرار گرفته است که دستیابی به روش‌ها و فنون جدید پیش‌بینی دقیق تر نوسان‌های قیمت‌ها در بورس از جمله آرمان‌های سرمایه‌گذاران و دست‌اندرکاران آن در سراسر دنیا است. در این تحقیق قابلیت پیش‌بینی شاخص بورس کل قیمت مورد تأیید قرار گرفته است.

علیرضا پاکدین امیری و همکاران (۱۳۸۸) به ارائه مدلی جهت پیش‌بینی شاخص کل قیمت سهام با رویکرد شبکه های عصبی پرداخته اند. در این پژوهش شاخص صنعت، شاخص مالی و شاخص بازده نقدی به صورت سالیانه به عنوان متغیرهای ورودی (مستقل) در نظر گرفته شده است. برای ارزیابی آن از مدل شبکه عصبی از طرح MLP با الگوریتم آموزش پس انتشار و مدل چند عاملی بهره گرفته شده است. نتایج نشان می دهد که مدل شبکه عصبی پیشنهادی، توانایی بالایی در پیش‌بینی شاخص قیمت سهام در بورس اوراق بهادار تهران را دارا می باشد.

سید علی نبوی چاشمی و همکاران (۱۳۹۰) به بررسی کارایی شاخص MA در تحلیل تکنیکال در پیش‌بینی قیمت سهام پرداخته اند. در این تحقیق از یکی از ساده ترین و در عین حال پرکاربردترین روش های تحلیل تکنیکی یعنی میانگین متحرک، میانگین متحرک وزنی و میانگین متحرک نمایی مورد استفاده قرار گرفته است. نتیجه بررسی نشان می دهد در مقایسه بین سه روش متحرک نمایی از لحاظ شاخص های اعتبار سنجی (میانگین قدر مطلق خطاها و نشانگر ردیاب) از اعتبار بالایی برای پیش‌بینی قیمت سهام برخوردار است و از قابلیت اطمینان بیشتری برخوردار می باشد.

اسدالله کردنائیج و همکاران (۱۳۹۲) به پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از انفیس پرداخته اند. در این تحقیق، ابتدا سه متغیر کلان اقتصادی به همراه مقادیر تاریخی تدپیکس به عنوان ورودی های مدل انتخاب شده اند؛ سپس ساختارهای گوناگون انفیس و شبکه عصبی مصنوعی پس انتشار خطا برای بررسی پیش‌بینی پذیری و شناسایی مدل انتخاب گردید. پیش‌بینی ها از طریق معیار غیر آماری نرخ برخورد، از نظر اقتصادی نیز بررسی شده اند. نتایج به دست آمده به کمک خطای آماری، غیر آماری و ضریب تعیین بیش از ۸۰ درصد نشان می دهد: مدل های ساده تر در پیش‌بینی شاخص، دارای دقت و عملکرد بیشتر هستند؛ و پیش‌بینی انفیس نسبت به شبکه های عصبی مصنوعی پس انتشار خطا

دقیق تر هستند. بر این اساس انفیس تکنیک امیدوار کننده ای برای سرمایه گذاران در پیش بینی شاخص بازده کل بورس اوراق بهادار تهران است.

عباسپور (۱۳۸۱) مطالعه ای جهت پیش بینی قیمت سهام شرکت "ایران خودرو" در بازار بورس تهران با استفاده از شبکه عصبی مصنوعی انجام داده و از داده های روزانه برای دوره زمانی ۱۳۷۹-۱۳۸۰ استفاده نمود. بر اساس یافته های تحقیق متغیر های موثر بر قیمت سهام شرکت "ایران خودرو" شامل نرخ ارز، قیمت نفت، نسبت P/E (قیمت به درآمد) و حجم مبادلات سهام بود. نتایج این تحقیق نشان از برتری نتایج حاصل از پیش بینی قیمت توسط شبکه عصبی مصنوعی نسبت به روش باکس-جنکنیز می باشد (عباس پور م.، ۱۳۸۱).

سینایی، مرتضوی و تیموری اصل دار (۱۳۸۴) به پیش بینی شاخص قیمت سهام در بورس اوراق بهادار تهران به وسیله شبکه عصبی مصنوعی و ارائه شواهدی مبنی بر رفتار آشوب گونه شاخص قیمت سهام در بورس اوراق بهادار پرداختند. آنها دو مجموعه از داده ها برای ورودی شبکه عصبی انتخاب نمودند، وقفه های مختلفی از شاخص و عوامل کلان اقتصادی به عنوان متغیر مستقل انتخاب کردند. در این تحقیق از مدل خطی ARIMA برای پیش بینی شاخص قیمت در هفته های بعدی استفاده شده است. نتایج حاصل از پژوهش نشان می دهد شبکه ی عصبی عملکرد بهتری نسبت به مدل خطی ARIMA برای پیش بینی شاخص قیمت دارد (سینایی، مرتضوی، & تیموری اصل، پیش بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه های عصبی مصنوعی، ۱۳۸۱).

مشیری و مروت (۱۳۸۵) شاخص کل بازده سهام را با مدل های خطی و غیر خطی پیش بینی نمودند. آنها با استفاده از داده های روزانه و هفتگی شاخص در بازه زمانی ۱۳۷۷ تا ۱۳۸۲ و به کار گیری از روشهای مختلف پیش بینی مانند مدل های GARCH و ARIMA و شبکه عصبی، شاخص کل را پیش بینی

نمودند. نتیجه حاکی از آن است که مدل شبکه عصبی، خطایی کمتر از دو مدل دیگر دارد. اما در آزمون معناداری مشخص شد که این تفاوت ها معنا دار نیستند به عبارت دیگر، دقت مدل های پیش بینی از نظر آماری تفاوت معناداری ندارد (مشیری & مروت، ۱۳۸۵).

همچنین عادل آذر و همکارانش در سال ۱۳۸۵ در پژوهشی پیش بینی شاخص سهام را با سه رویکرد روش های کلاسیک، رویکرد هوش مصنوعی و رویکرد ترکیبی انجام دادند. نتایج این تحقیق بیانگر این حقیقت است که شبکه های عصبی فازی بر روش ARIMA برتری داشته و دارای ویژگی های منحصر بفرد همگرایی سریع و دقت بالا هستند و برای پیش بینی شاخص قیمت سهام مناسب می باشند (آذر & افسر، ۱۳۸۵).

منجمی و همکاران (۱۳۸۸) در پژوهشی تحت عنوان "پیش بینی قیمت سهام در بازار بورس اوراق بهادار با استفاده از شبکه های عصبی- فازی و الگوریتم های ژنتیک و مقایسه ی آن با شبکه های عصبی مصنوعی" نشان دادند که از نقطه نظر معیارهای ارزیابی عملکرد، پیش بینی قیمت سهام روز بعد توسط مدل ترکیبی شبکه های عصبی- فازی و الگوریتم ژنتیکی دقیقتر از شبکه های عصبی است. به عبارتی دیگر، پیش بینی قیمت سهام با استفاده از شبکه های عصبی - فازی و الگوریتم های ژنتیکی، خطای برآورد قیمت سهام را نسبت به تکنیک شبکه های عصبی مصنوعی کاهش میدهد (منجمی، ابزری، & رعیتی شوارزی، ۱۳۸۸).

۲-۱۱-۲- پیشینه خارجی

فیسارن سودی بانجراد^۱ و همکاران (۲۰۱۰) پژوهشی در زمینه پیش‌بینی بازار اوراق بهادار تایلند با استفاده از استراتژی‌های تکاملی انجام دادند. در این مطالعه یک تابع پیش‌بینی برای بازار بورس و اوراق بهادار تایلند (SET^۲ index) پیشنهاد شده است. در تابع پیش‌بینی عوامل مهم اقتصادی از قبیل داو جونز^۳، نیکی^۴ و شاخص‌های هنگ سنگ^۵ و شاخص قبلی بازار بورس اوراق بهادار تایلند شامل می‌شود.

ایم^۶ در سال ۲۰۰۲ با استفاده از مدل شبکه عصبی، بازدهی شاخص روزانه سهام برزیل را پیش‌بینی نمود. وی در مدل سازی خود از داده‌های روزانه شاخص سهام برزیل طی سالهای ۱۹۹۴ تا ۱۹۹۸ استفاده نمود و نتایج پیش‌بینی را با استفاده از معیارهای ریشه میانگین مجذور خطا (RMSE) و میانگین قدر مطلق خطا (MAE) با نتایج پیش‌بینی مدل‌های GARCH و ARIMA مقایسه نمود و برتری شبکه عصبی مصنوعی را نشان داد (ایم، ۲۰۰۲).

ژی یانگ ژانگ^۷ (۲۰۰۶) به پیش‌بینی روند قیمت سهام بورس شانگ‌های با استفاده از ماشین بردار پشتیبان (SVM) پرداخت. او شاخص روزانه قیمت سهام بازار شانگ‌های را از سال ۲۰۰۳ تا ۲۰۰۵ استخراج نمود و آنها را به دو دسته آموزش و آزمون تقسیم نمود. همچنین از توصیه‌های تقریباً ۴۰۰ تحلیلگر بازار سرمایه و پیش‌بینی آنها به عنوان متغیر ورودی استفاده نمود. نتیجه مطالعه او نشان داد ماشین بردار پشتیبان (SVM) قابلیت پیش‌بینی بالایی دارد و ترکیب ماشین بردار پشتیبان (SVM) با مدل‌های هوشمند حتی نتیجه بهتری از خود مدل ماشین بردار پشتیبان (SVM) دارد (ژانگ، ۲۰۰۵).

¹ Phaisarn Sutheebanjard

² Stock Exchange of Thailand

³ Dow Jones

⁴ Nikkei

⁵ Hang Seng

⁶ Yim

⁷ Zhang et al

ام - تی سانگ و همکارانش^۱ در سال ۲۰۰۷ کاربرد شبکه عصبی NN5^۲ را در پیش‌بینی قیمت سهام هنگ کنگ بررسی نمودند. این سیستم بر روی داده های سهام دو شرکت سهامی بانکداری هنگ کنگ و شانگ های آزمون شده است. این سیستم نرخ موفقیت کلی بیش از ۷۰ درصد را نشان می دهد (تی سانگ، ۲۰۰۷).

کلی لوگان^۳ در سال ۲۰۰۷ به پیش‌بینی میزان حجم پول در اقتصاد آمریکا توسط مدل تخمینگر حداقل درجه و روش های بیزی پرداخته است. او از متغیرهای نرخ بهره بلندمدت، نرخ بهره کوتاه مدت، نرخ بیکاری، میزان سپرده گذاری و هزینه خدمات پولی در بازه زمانی ۱۹۶۰ تا ۲۰۰۹ به صورت ماهیانه استفاده کرده است. نتایج تحقیق نشان از برتری پیش‌بینی مدل تخمینگر حداقل درجه می دهد (چی لی، ۲۰۰۷).

مینگ چی لی^۴ (۲۰۰۹) به پیش‌بینی شاخص نزدک NASDAQ با یک مدل ترکیبی تخمینگر بردار پشتیبان (SVR) و مقایسه آن با شبکه های عصبی پرداخته است. در این پژوهش مدل تخمینگر بردار پشتیبان (SVR) با تابع FSSFS^۵ ترکیب و توسط ۲۹ شاخص فنی به عنوان مجموعه ای از ویژگی های کامل در جهت تغییر شاخص استفاده گردیده است. داده های تحقیق از سال ۲۰۰۱ تا ۲۰۰۷ بوده که ۸۰٪ داده ها جهت آموزش مدل و ۲۰٪ داده ها برای آزمون استفاده گردیده است. نتایج تحقیق نشان از برتری مدل ترکیبی تخمینگر بردار پشتیبان (SVR) نسبت به شبکه عصبی می باشد (چی لی^۶، ۲۰۰۹).

¹ M.Tsang et al

² Neural Network 5

³ Kelly Logan

⁴ Ming-Chi Lee

⁵ F-score and Supported Sequential Forward Search

⁶ Chi Lee

یاکوب کارا و همکاران^۱ (۲۰۱۱) در پژوهشی به پیش‌بینی جهت حرکت شاخص قیمت سهام بورس استانبول با مدل‌های شبکه عصبی - فازی و ماشین بردار پشتیبان (SVM) پرداختند و از داده‌های روزانه ۱۹۹۷ تا ۲۰۰۷ به همراه ۱۰ شاخص فنی به عنوان متغیرهای ورودی مدل استفاده گردیده است. شبکه عصبی - فازی ۷۴/۷۵٪ و مدل ماشین بردار پشتیبان ۵۲/۷۱٪ از عهده پیش‌بینی برآمدند و عملکرد بهتر شبکه عصبی - فازی نسبت به مدل ماشین بردار پشتیبان حاصل گردید. همچنین بهترین عملکرد پیش‌بینی متعلق به سال ۲۰۰۱ بوده است (یاکوب^۲ و همکاران، ۲۰۱۱).

¹ Yakup Kara et al

² Yakub

۳- فصل سوم

روش شناسی پژوهش

الگوریتم های تکاملی به طور موفقیت آمیزی برای تولید خودکار برنامه ها مورد استفاده قرار می گیرد. به طور خاص برنامه نویسی ژنتیک محبوبیت قابل توجه برخوردار است و دارای استفاده گسترده در زمینه های مدل سازی اقتصادی، سیستم های بازرگانی، خرید و فروش ارز، تضمین وام، پردازش تصویر، تحقیقات زیست شناسی و بیولوژیکی، کنترل و هدایت ربات و ... است. در سال های اخیر از الگوریتم تکاملی و دیگر فرامکاشفه های زیستی، برای حل مسائل گوناگون جستجو و بهینه سازی های پیوسته و گسسته به طور شگفت آوری رشد فزاینده ای داشته است. مهمترین عامل استفاده از این نوع الگوریتم سادگی پیاده سازی و کاربرد آنها برای مسائل با ماهیت متفاوت است. اصولاً GP از زبان برنامه نویسی Lisp به عنوان زبان هدف استفاده می کند، اما در GE به هر زبان دلخواه می توان برنامه نویسی را انجام داد. بر خلاف GP تکامل دستوری نمی تواند روند تکاملی را روی برنامه های واقعی اجرا کند بلکه رشته های دو دویی با طول متغیر این کار را انجام می دهد. فرایند نگاشت به هر زبانی برای تولید برنامه ها با استفاده از رشته های دو دویی برای انتخاب قوانین تولید در فرم گرامر BNF بکار می رود. در نهایت ساختار یک برنامه از لحاظ دستوری صحیح از رشته های باینری که می تواند یک تابع سازگار را مورد ارزیابی قرار دهد، است (هارپر^۱ ۲۰۱۰).

همان طور که تاریخ الگوریتم های تکاملی نشان می دهد، گونه های زیادی از الگوریتم های تکاملی وجود دارند. ولی ایده همه آنها یکی است: با داشتن جمعیتی از گونه ها^۲، فشار محیطی باعث انتخاب می شود (القاء بهترین^۳) و این افزایش شایستگی^۴ جمعیت را نتیجه می دهد. با داشتن یک تابع کیفیتی که می خواهیم بیشینه شود، می توان مجموعه ای از جواب های کاندید را به طور تصادفی تولید کرد و تابع

¹ Harper

² individual

³ Survival of the fittest

⁴ fitness

کیفیت را به عنوان معیاری برای محاسبه شایستگی به کار برد - (هر چه بیشتر، بهتر) بر اساس این شایستگی، بعضی از کاندیدهای بهتر انتخاب می شوند، تا به عنوان هسته ای برای تولید نسل بعد به کار روند. بر روی این کاندیدها ترکیب و یا جهش^۱ اعمال می شود. ترکیب بر روی دو یا بیشتر کاندید اعمال می شود (والدین) و نتیجه آن تولید فرزند (فرزندانی) است (اونیل و رایان، ۲۰۰۱).

اعمال ترکیب و جهش باعث تولید مجموعه جدیدی می شود که با مجموعه قبلی (والدین) رقابت می کنند تا در نهایت برنده ها در نسل بعدی ظاهر شوند. این کار می تواند ادامه پیدا کند تا یک کاندید با ویژگی های کافی (جواب) به دست بیاید و یا اینکه محدودیت‌هایی که از قبل برای مسئله تعریف کرده ایم، ارضا شوند.

در این عمل دو نیروی اصلی وجود دارد که پایه سیستم تکاملی است:

- **عملگرهای تغییر (ترکیب و جهش)** که باعث ایجاد گوناگونی لازم و در نتیجه نوآوری می شود.

- **انتخاب** که نیرویی است که کیفیت را به جلو می برد.

ترکیب تغییر و انتخاب باعث بهتر شدن مقادیر شایستگی در جمعیت ها می شود.

با مشاهده روند حرکت جمعیت می توان تکامل به سوی بهینگی را مشاهده کرد.

تکامل به عنوان فرایند تطبیق بیان می شود. از این دید، شایستگی به عنوان هدف اصلی که باید بهینه شود مطرح نیست، بلکه عبارتی است که نیازمندی کل محیط را بیان می کند، هرچه این نیازمندی ها بیشتر ارضا شوند، در نتیجه تعداد بیشتری از اعضای جمعیت خود را نشان می دهد. عمل تکامل باعث می شود که جمعیت با محیط خود بیشتر و بیشتر سازگار شود (اونیل و رایان، ۲۰۰۱).

^۱ mutation

بسیاری از اجزای فرآیند تکامل اتفاقی^۱ هستند. این اجزا در زمان انتخاب موجوداتی که مناسب تر^۲ هستند، احتمال انتخاب بیشتری دارند، هر چند در بیشتر اوقات، موجودات ضعیف تر هم شانس انتخاب شدن و زنده ماندن را دارند. اکثر اوقات موجودات به طور تصادفی برای ترکیب از جمعیت خارج می شوند. این مطلب در مورد تغییرات نیز صادق است. طرح کلی الگوریتم های تکاملی در شکل ۱-۳ آمده است.

1. A current population of μ individuals is randomly initialized.
2. Fitness scores are assigned to each of the μ individuals.
3. λ new offspring are generated by recombination from the current population.
4. The λ new offspring are mutated.
5. Fitness scores are assigned to the λ new offspring.
6. A new population of μ individuals is selected, using either (μ, λ) or $(\mu + \lambda)$ selection.
7. The new population becomes the current population.
8. If the termination conditions are satisfied exit, otherwise go to step 3.

شکل ۱-۳: طرح کلی الگوریتم تکاملی

همان گونه که از شبه کد نیز معلوم است، الگوریتم های تکاملی جزئی از الگوریتم های تولید - آزمایش^۳ هستند.

- الگوریتم تکاملی مبتنی بر جمعیت است.
 - الگوریتم تکاملی از ترکیب استفاده می کند تا اطلاعات گونه های بیشتری را در یک گونه خلاصه کند.
 - الگوریتم تکاملی اتفاقی است.
- گونه های مختلف الگوریتم های تکاملی همگی از طرح کلی که ارائه شد، پیروی می کنند و فقط در جزئیات تکنیکی متفاوت هستند (ایبن^۱ و اسمیت^۲، ۲۰۱۳).

¹ stochastic

² Fitter

³ Generate-and-test

۳-۲-انواع الگوریتم های تکاملی

همان طور که در بخش اول گفته شد، گونه های مختلفی از الگوریتم های تکاملی وجود دارد، گفته می شود که این الگوریتم ها به سه دسته کلی تقسیم می شوند:

۱. الگوریتم های ژنتیکی (GA) ارایه شده توسط Holland (هالند، ۲۰۰۲) و مطالعه شده توسط Goldberg (گلدبرگ، ۲۰۰۸)

۲. استراتژی های تکاملی (ES) ارایه شده توسط Rechenberg (ریچنبرگ، ۱۹۷۳).

۳. برنامه ریزی تکاملی (EP) ارائه شده توسط L.J. Fogel et. Al (فوگل، ۱۹۶۶). و اصلاح شده توسط فوگل.

هر کدام سه روش بالا اثبات شده اند که با داشتن فضای مسئله پیچیده، پیوسته و چند کیفیت^۳ به جواب تقریباً بهینه می رسند.

از دسته بندی بالا، گروه اول (الگوریتم های ژنتیکی) را به تفسیر در فصول بعدی بررسی می کنیم. در مورد دو دسته دیگر به طور اختصار توضیح می دهیم.

۳-۳-استراتژی های تکاملی

استراتژی های تکاملی ارایه شده توسط ریچنبرگ از لحاظ تاریخی به منظور حل مسائل بهینه سازی پارامترها تعریف شدند. در نتیجه هر گونه در جمعیت به صورت لیستی از اعداد حقیقی تعریف می شد.^۴ علاوه بر این هر گونه شامل یک سری پارامترهای استراتژی نیز بود. این پارامترها برای کنترل رفتار عملگرهای جهش استفاده می شدند.

¹ Eiben

² Smith

³ Multi modal

^۴ به این بخش از گونه متغیرهای شی (object variables) می گویند.

در هر دوره از اجرای الگوریتم، λ فرزند تولید می شوند. اندازه جمعیت μ بود و معمولاً λ برابر μ ۶ بود عملگر ترکیب به ازای تولید یک فرزند نیاز به دو والد داشت. همین دو والد برای تولید هر دو بخش فرزند (پارامترهای استراتژیک و متغیرهای شی) مورد استفاده قرار می گرفتند. انتخاب والدا به صورت تصادفی از جمعیت فعلی بود. گونه های مختلفی برای ترکیب وجود دارند ولی بهترین جواب برای نوع ترکیبی بود که متغیرهای شیء فرزند را همه از روی یکی از والدین برداشته و برای پارامترهای استراتژی نیز میانگین پارامترهای استراتژی والدین استفاده شود.

عملگر اصلی در ES، جهش است و روی هر دو بخش گونه اعمال می شود. عملگر جهش ابتدا روی بخش پارامترهای استراتژی اعمال می شود. بعد از آن متغیرهای شیء با استفاده از نتایجی که از اعمال جهش روی پارامترهای استراتژی حاصل شده است، جهش می کنند. این نوع جهش به ES این قابلیت را می دهد که خود را تطبیق دهد و سعی کند بهترین پارامترهای استراتژی را بدست آورد.

انتخاب در ES به طور قطعی انجام می شود. بهترین μ گونه از بین λ فرزند انتخاب می شوند. (انتخاب (μ, λ)) یا بهترین μ گونه از اجتماع دو جمعیت انتخاب می شوند (انتخاب $(\mu + \lambda)$) همانند برخی دیگر از الگوریتم های تکاملی ES تا زمانی که یکی از شروط خاتمه مسئله برقرار شود، اجرا می شود. در شکل ۱-۲ شبه کدی از ES ارائه شده است.

۳-۴- برنامه ریزی تکاملی

برنامه ریزی تکاملی ارائه شده توسط فوگل اولین بار برای تکامل ماشین های حالت با استفاده از تعداد محدودی کد کردن الفباها تعریف شد. بعد از آن فوگل را برای بهینه سازی متغیرها گسترش داد. همانند ES، گونه ها در EP به صورت رشته ای از اعداد حقیقی تعریف می شوند. تفاوت EP با ES این

است که در EP عملگر ترکیب وجود ندارد. تکامل به طور کامل به عملگر جهش وابسته است. عملگر جهش از توزیع احتمال گووس برای تغییر هر متغیر استفاده می کند.

انحراف معیار استاندارد از روی ریشه تبدیل خطی شایستگی والدین به دست می آید (این تبدیل را باید پارامتری کنیم) برای جلوگیری از کار زیاد با پارامترها فوگل meta-EP را ارائه کرده است. در meta-EP هر گونه شامل متغیرهای شی و واریانس (یک واریانس به ازای هر متغیر شی) است. از این واریانس ها برای کنترل عملگر جهش گووسی استفاده می شود.

عملگر انتخاب با استفاده از tournament-select^۱ انجام می شوند. گونه خاصی به نام q-tournament selection استفاده می شود. الگوریتم به این صورت است که u را اجتماع جمعیت فرزندان و والدین می گیریم. به ازای هر عضو m از q, u حریف از u به صورت تصادفی انتخاب می شوند. تعداد حریف هایی که مقدار شایستگی آنها از m بدتر است حساب می شود.^۲ μ گونه ای که تعداد حریف های ضعیف آنها از همه بیشتر هستند و وارد جمعیت می شوند. هرچه q افزایش پیدا می کند، الگوریتم انتخاب قطعی تر می شود. در نتیجه بهترین گونه همیشه در جمعیت جدید حضور دارد. شبه کد EP در شکل ۱-۳ ارائه شده است (جونز^۳، ۲۰۰۱)

^۱ انتخاب ترنمنت

^۲ Count

^۳ G.Jones

1. A current population of μ individuals is randomly initialized.
2. Fitness scores are assigned to each of the μ individuals.
3. The mutation operator is applied to each of the μ individuals in the current population to produce μ offspring.
4. Fitness scores are assigned to the μ offspring.
5. A new population of size μ is created from the μ parents and the μ offspring using tournament selection.
6. If the termination conditions are satisfied exit, otherwise go to step 3.

شکل ۳-۲: شبه کد برنامه ریزی تکاملی

۳-۵- الگوریتم ژنتیک

محدوده کاری الگوریتم ژنتیک بسیار وسیع می باشد و هر روز با پیشرفت روزافزون علوم و تکنولوژی استفاده از این روش در بهینه سازی و حل مسائل بسیار گسترش یافته است. الگوریتم ژنتیک یکی از زیر مجموعه های محاسبات تکامل یافته می باشد که رابطه مستقیمی با مبحث هوش مصنوعی دارد در واقع الگوریتم ژنتیک یکی از زیر مجموعه های هوش مصنوعی می باشد. الگوریتم ژنتیک را می توان یک روش جستجوی کلی نامید که از قوانین تکامل بیولوژیک طبیعی تقلید می کند. الگوریتم ژنتیک بر روی یکسری از جواب های مساله به امید بدست آوردن جوابهای بهتر قانون بقای بهترین را اعمال می کند. در هر نسل به کمک فرآیند انتخابی متناسب با ارزش جواب ها و تولید مثل جواب های انتخاب شده به کمک عملگرهایی که از ژنتیک طبیعی تقلید شده اند، تقریب های بهتری از جواب نهایی بدست می آید. این فرایند باعث می شود که نسلهای جدید با شرایط مساله سازگارتر باشد (گلدبرگ، ۱۹۸۹).

۳-۶- الگوریتم تکامل دستوری

الگوریتم تکامل گرامری یک شاخه جدید از الگوریتم های تکاملی که توسط رایان، کولینز، اونیل و باربازون پیشنهاد شده است. GE بر حسب عبارات جبری، قوانین تصمیم گیری، قوانین دسته بندی شکل گرفته است. این روش تکاملی از برنامه ژنتیکی بر گرفته شده است. اغلب برای یادگیری سیستم های حمایت تصمیم گیری، سیستم های تخصصی سیستم های دسته بندی بکار برده می شود (اونیل و باربازون، ۲۰۰۴). با توجه به کاستی های GP سنتی و رویکرد GA و به ویژگی های مطلوب مشخص شده زیر، GE را می توان به عنوان وسیله ای در نظر گرفت که برخی از مشکلات را حل خواهد کرد. تکامل دستوری را می توان به عنوان تکنیکی در نظر گرفت که کانر رایان، کالینز و مایکل اونیل در دانشگاه لیمرک^۱ کشور ایرلند ابداع کننده آن هستند. تحقیقات نشان داده است که GE بر مسائل از قبیل رگرسیون نمادین (اونیل و رایان، تکامل دستوری، ۲۰۰۳)، رتبه بندی اعتباری برای اوراق قرضه (باربازون و اونیل، نرخ اعتبار صادر کننده اوراق قرضه با تکامل دستوری، ۲۰۰۴)، شاخص بازار سهام (اونیل و همکاران، قوانین شاخص بازار تجاری با استفاده از تکامل دستوری، ۲۰۰۱) و تشخیص لبه برای پردازش تصویر (دریسکال & همکاران، ۲۰۰۲) اثربخش بوده است. GE یک راه منحصربفردی با استفاده از گرامر در فرایند برنامه نویسی اتوماتیک ارائه می دهد. رشته های باینری با طول متغیر مورد استفاده قرار می گیرد، هر کد نشان دهنده یک مقدار عدد صحیح که این کدها منبع ۸ بیتی طبق باز تولید کد ژنتیک ساخته شده است. مقادیر ارزش در یک تابع نگاشت برای انتخاب قوانین تولید مناسب از BNF تعریف شده انتخاب می شود؛ تعداد تولید شده معمولاً نشان دهنده یکی از قوانینی است که باید بکار گرفته شود. GE از مشکل اعتنا به مقادیر عدد صحیح رنج نمی برد چون GE ارزش غیر قانونی تولید نمی کند. موضوع حصول اطمینان از نقشه کامل یک فرد بر روی یک برنامه که بطور انحصاری شامل پایانه هایی با استفاده از تکنیک های

¹ University of Limerick

جدید الگوریتم تکاملی به نام بسته بندی حل می‌شود. این روش با الهام از تداخل ژن های نمایش گذاشته شده توسط بسیاری از باکتری ها، ویروس ها و میتوکاندریهای که توانایی به استفاده مجدد از مواد ژنتیکی مشابه در شکل ژن های مختلف دارند، است (اونیل، ۲۰۱۰). GE پس از آن یک سیستمی است که یک فرآیند نگاشت جدید قوی بکار می‌گیرد، در پایان نتایج توانایی تولید کد با هر زبانی از رشته باینری های ساده را دارد. در حال حاضر عنصر جستجوی سیستم توسط یک الگوریتم تکاملی انجام می‌شود، اگرچه هر روش تحقیق امکان پذیری با توانایی بیش از طول متغیر رشته های باینری می‌تواند بکار بگیرد. به طور خاص، پیشرفت های آینده در زمینه الگوریتم های تکاملی می‌تواند به راحتی به این سیستم ارائه برنامه گنجانده شود.

۳-۷-گرامر

آموزش گرامر یکی از بخش های مهم در حل مسئله با استفاده از GE است. تغییرات کوچک می‌تواند تأثیرات بسیار زیادی بر روی تحقیق بگذارد. یک گرامر رسمی یک وسیله مخصوص زبان نحوی، شامل مجموعه قوانین که مقرر می‌کند چطور یک رشته ای ایجاد کند و شروع یک نماد باشد. چامسکی^۱ یک سلسله از زبان به همراه گرامر که به با متن های زبان و دستور آزاد قابل استفاده در GE بود، توصیف و تعریف کرد. گرامر اغلب در زمینه محاسبات و توصیف نحو زبان های برنامه بکار می‌رود. در اینجا از چهار جز (N, T, P, S) جهت توصیف گرامر بکار می‌رود. اجزا مجموعه ای از غیرترمینال ها (N)، مجموعه ای از ترمینال ها (T)، مجموعه ای از قوانین تولید (P) و نماد شروع (S) است. غیر ترمینال ها نمادهایی هستند که به دیگر غیر ترمینال ها یا ترمینال ها توسط گرامر ارجاع داده می‌شوند. ترمینال ها نمادهای واقعی هستند که شامل زبان هایی که با گرامر توصیف می‌شوند و با قوانین تولید به هیچ چیز دیگری تبدیل نمی‌شوند. قوانین تولید تبدیل یک نماد (غیر ترمینال ها) به دیگر نمادها را تعریف می‌کند. یک

¹ Chamesky

نماد شروع یک نشانه و سیگنالی است که به عنوان اولین غیر ترمینال، که ممکن است با بکار گیری قوانین تولید صورت پذیرد، بکار گرفته می‌شود. استاندارد علامت گذاری بکار رفته در این الگوریتم جهت توصیف یک گرامر رسمی^۱ BNF شناخته می‌شود. BNF شامل یک سری از قوانین تولید با غیر ترمینال ها در سمت چپ و یک ترمینال یا دیگر غیر ترمینال ها در سمت راست در شکل ۳-۱ در زیر آورده شده است. نماد شروع اولین غیر ترمینال ها نشان داده شده است (کولینز^۲ و ریان^۳، ۱۹۹۸).

```

<expr>      ::= <expr><op><expr>
              | (<expr><op><expr>)
              | <const>
              | <var>

<op>        ::= +
              | -
              | *
              | /

<var>       ::= X

<const>     ::= 1
              | 2
              | 3
              | 4

```

شکل ۳-۳: یک نمونه از گرامر در فرم BNF

غیر ترمینال <op> در نمونه گرامر می‌تواند با یکی از چهار ترمینال ها مربوط به عملگرهای محاسباتی جمع، تفریق، ضرب و تقسیم جایگزین شود.

مثال: یک گرامر برای تولید عبارات منطقی از قبیل عبارات زیر استفاده می‌کند.

((true or false) and not true):

- Non-terminals:

$N = \{ \langle \text{expr} \rangle, \langle \text{biop} \rangle, \langle \text{uop} \rangle, \langle \text{bool} \rangle \}$

- Terminals:

¹ Backus Naur Form

² Collins

³ Ryan

$T = \{ \text{and, or, xor, nand, not, true, false, (,)} \}$

- Strat symbol:

$S = \{ \langle \text{expr} \rangle \}$

- Production rules:

$\langle \text{expr} \rangle ::= (\langle \text{expr} \rangle \langle \text{biop} \rangle \langle \text{expr} \rangle) \quad (0)$

$\quad \quad \quad I \langle \text{uop} \rangle \langle \text{expr} \rangle \quad (1)$

$\quad \quad \quad I \langle \text{bool} \rangle \quad (2)$

$\langle \text{biop} \rangle ::= \text{and} \quad (0)$

$\quad \quad \quad I \text{ or} \quad (1)$

$\quad \quad \quad I \text{ xor} \quad (2)$

$\langle \text{uop} \rangle ::= \text{not} \quad (0)$

$\langle \text{bool} \rangle ::= \text{true} \quad (0)$

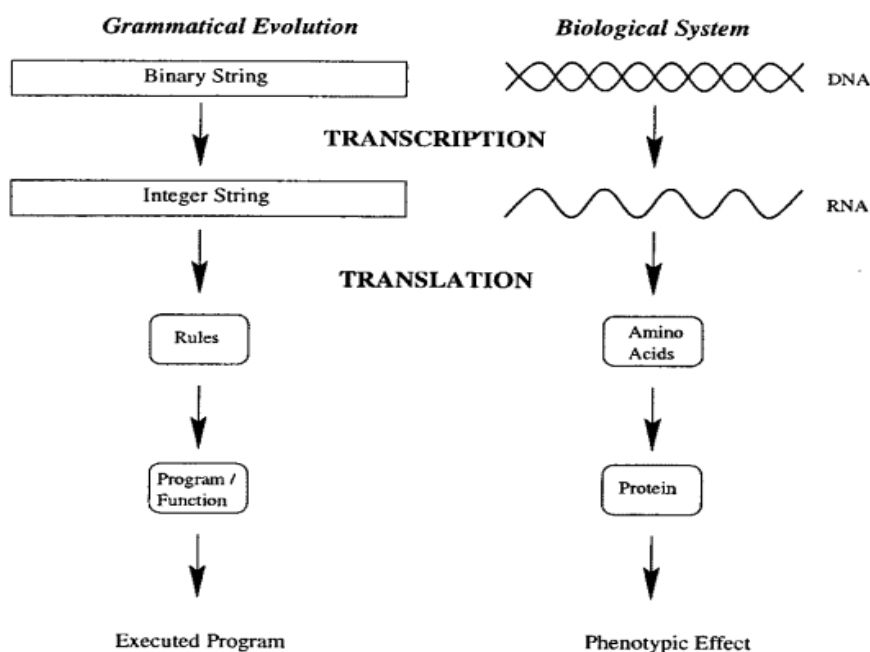
$\quad \quad \quad I \text{ false} \quad (1)$

۳-۷-۱-رمزگشایی

فنونتیپ یک عبارت صحیح در تعریف گرامر است. یک عبارت که از نماد شروع توسط نماد های ترکیبی غیر ترمینال ها و ترمینال ها بر طبق قوانین تولید جایگزین می شود، می توان بعنوان مثال ذکر کرد. ژنونتیپ یک نمایش از دوتایی از تعریف متوالی اعداد صحیح است هر کد با یک نماد غیر ترمینال مطابقت دارد و تبدیل آن با قوانین تولید تعیین می کند.

۳-۸- تطابق GE و سیستم بیولوژیک

شکل زیر نشان دهنده ی مقایسه ی بین سیستم GE و یک سیستم ژنتیکی بیولوژیکی است.



شکل ۳-۴: تطابق سیستم GE و سیستم بیولوژیکی تولید پروتئین

رشته باینری مربوط به ژنوتیپ یک فرد در GE معادل مارپیچ گونه DNA یک موجد زنده است که هر کدام تشکیل فنوتیپ را هدایت می کنند. در خصوص GE این از طریق بکار گیری قوانین تولید برای ایجاد ترمینال های این برنامه رخ می دهد. در مورد پروتئین فنوتیپی را با تعیین ترتیب و نوع زیر مجموعه پروتئین مثل آمینو اسید ها که به هم پیوسته اند، به تولید پروتئین می انجامد.

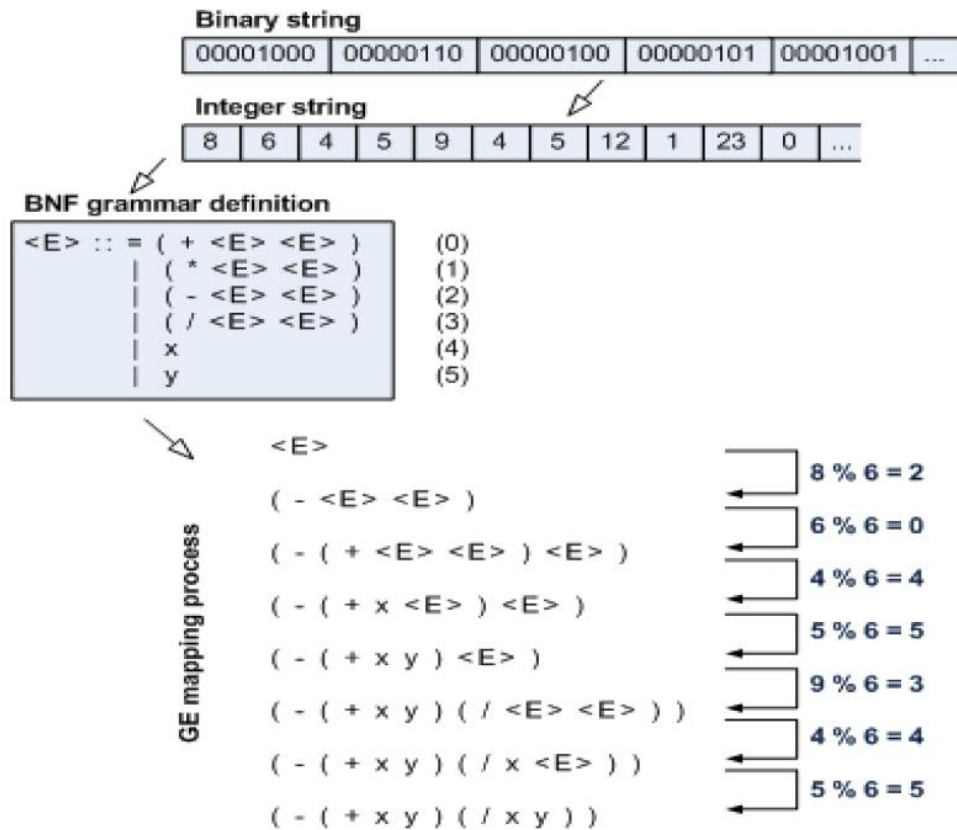
مراحل زیر در تکامل گرامری انجام می شود.

۱. ژنوتیپ (رشته باینری با طول متغیر) برای نگاشت نماد شروع تعریف گرامری BNF به ترمینال ها، مورد استفاده قرار می گیرد. گرامر برای مشخص کردن فنوتیپ های قانونی استفاده می شود.
۲. کد الگوریتم GE معمولا هشت بیتی هستند و عدد صحیح متناظر با ترتیب کد برای تعیین کدام قانون برای غیر ترمینال ها تفسیر شود و بکار رود، فراخوانی می کند و از آن استفاده می کند. اگر وقتی که کد ها را می خواند، به انتهای ژنوتیپ برسد، دوباره از اول ژنوتیپ شروع می کند.

۳. شکل قانون تولید با استفاده از فرمول $\text{form} = \text{codon mod forms}$ محاسبه شود که در آن

codon و forms تعداد شکل های جایگزین برای غیر ترمینال فعلی است.

یک مثال از فرآیند نگاشت که با تکامل گرامری به کار گرفته شده است در شکل زیر آمده است.



شکل ۳-۵: فرآیند نگاشت GE

در این مثال، کد اول مربوط به ژنوتیپ فرد، رشته دودویی ۰۰۰۰۱۰۰۰ است که شکل باینری عدد صحیح ۸ است. سمبل شروع <E> شش شکل جایگزین دارد. بنابراین شکلی که باید به کار گرفته شود اینار با برچسب ۸/۶ یعنی ۲ است (باقیمانده تقسیم عدد ۸ بر ۶ که برابر ۲ است)، که از عبارت حاصل می‌شود. سپس به منظور جایگزین کردن سمبل غیر ترمینال اول <E> مربوط به عبارت جدید، کد بعدی خوانده می‌شود و این تا زمانی که عبارت تنها شامل سمبل های ترمینال X و Y و هر کدام از عملگرهای

محاسباتی باشد، ادامه پیدا می کند. یعنی تا زمانی که تمامی نشانه های غیر ترمینالی، جایگزین شوند. پس از فرایند نگاشت (ایجاد فنوتیپ)، امتیاز تناسب محاسبه شده و بر طبق مشخصات مسئله مشخص به هر فرد (فنوتیپ) اختصاص داده می شود. این امتیازهای تناسب به الگوریتم تکاملی باز ارسال می شود که از آنها به منظور تکامل یک توده از افراد جدید استفاده می کند.

۳-۹-نگاشت ژنو تیپ^۱ به فنو تیپ^۲

GE یک الگوریتم تکاملی است که می تواند ژنو تیپ را از فنو تیپ جدا کند، یک روش که شبیه فرآیند های بیولوژیکی همانند تقلید^۳ EA از GA^۴ یا GP^۵ است. در تبدیل ژنو تیپ به فنو تیپ که شبیه ژن سیستم بیولوژیکی است، با اجرای قوانین تولید از یک گرامر و دستور زبان آزاد صورت گرفته است. این گرامر در فرم BNF و برای ژنو تیپ های فردی که شامل جمعیت هم می شود، تعریف می شود. فنو تیپ با استفاده از یک تابع سازگار بجای ژنو تیپ ارزیابی را انجام می دهد. اساس منحصر به فرد در GE، با ژنو تیپ که معمولا با یک بردار عدد صحیح صورت می گیرد، مشخص می شود. هر فرد می تواند بعد از نگاشت ژنو تیپ به فنو تیپ با یک دسته زبان خاص همراه باشد. این فرایند با شروع در اولین کد و یا عدد صحیح در ژنو تیپ اجرا می شود و برای تولید قوانین تولید از دستور زبان استفاده می شود. اساس هر نفر در GE با فنو تیپش، بطور کلی با بردار عدد صحیح نمایش داده می شود. هر فرد بعد از قرار گرفتن در نگاشت ژنو تیپ به فنو تیپ ارزیابی شود. این فرایند با شروع در اولین کد یا عدد صحیح، در ژنوتیپ و استفاده در انتخاب قانون تولید از گرامر اجرا می شود. هر غیر ترمینال در سمت چپ می تواند به یک یا بیشتر از یک مقدار (ترمینال یا غیر ترمینال) در سمت راست تبدیل شود. برای هر غیر ترمینال در سمت چپ N قانون وجود دارد که می تواند از ۰ تا N-1 شماره با مقدار کدی که برای ساخت فنو تیپ هر عضو است، مورد

¹ Genotype

² Phenotype

³ Evolution Algorithm

⁴ Genetic Algorithm

⁵ Genetic Programming

استفاده قرار گیرد. معیار الگوریتم تکامل دستوری با بکار گیری عملگرهای مدل برای کد، با بهره گیری از شماره های ۰ تا N-1، یک قانون تولید انتخاب می کند. این تعداد برای تولید قانون برای بکار گیری غیر ترمینال ها انتخاب می شوند. در اولین کد یک قانون تولید انتخاب و به اولین نماد شروع تبدیل می شود. کد بعدی برای تولید قانون دوباره، از باقی مانده غیر ترمینال ها دوباره همین فرایند تکرار می شود. ای فرایند تا زمانی ادامه پیدا می کند که هیچ غیر ترمینال باقی نماند و به جواب لازم رسیده باشیم.

- A) $\langle \text{expr} \rangle ::= \langle \text{expr} \rangle \langle \text{op} \rangle \langle \text{expr} \rangle$ (0)
 | $(\langle \text{expr} \rangle \langle \text{op} \rangle \langle \text{expr} \rangle)$ (1)
 | $\langle \text{const} \rangle$ (2)
 | $\langle \text{var} \rangle$ (3)
- B) $\langle \text{op} \rangle ::= +$ (0)
 | $-$ (1)
 | $*$ (2)
 | $/$ (3)
- C) $\langle \text{var} \rangle ::= X$ (0)
- D) $\langle \text{const} \rangle ::= 1$ (0)
 | 2 (1)
 | 3 (2)
 | 4 (3)

شکل ۳-۶: یک گرامر علامت گذاری شده که قوانین تولید با شاخص های مرتبط در آن نشان می دهد.

۳-۱۰-تلفیق در GE

ماهیت نگاشت از ژنوتیپ به فنو تیپ در GE بدان معناست که کد داده یا مجموعه کدها در ژنوتیپ هیچ معنای ذاتی ندارد و تنها در چارچوب یک نگاشت کامل تفسیر شده است. بنابراین یک گروه از کدون ها که برای یک رشته داده شده از اولیا می تواند به خوبی به یک رشته مختلف کامل، وقتی که به یک کودک تغییر می دهد، کد بوجود بیاید. فنو تیپ حاصل از ایجاد یک فرد جدید با استفاده از موجودار شدن ادغام ممکن است شباهت کمی به هر یک از اولیا ، بسته به گرامر، نقطه برش و کد مخصوصی که تشکیل شده

است. بانژاف^۱، کنرادس^۲، فرانکونه^۳ و نوردین^۴ یک نوع شناخته شده از ادغام همولوگ^۵ که به دنبال مدل باز ترکیبی که در موجودات اتفاق می افتد را گسترش دادند (بانژاف و همکاران، ۱۹۹۹).

به طور خاص این رویکرد تلاش می کند به چنین نقاط ادغامی اجازه می دهد یک بخش از یکی از والدین در همان موقعیت در یک اولیای دیگر تعویض شود. این شبیه به مدلی است که ادغام بیولوژیکی با DNA از هر دو اولیا ساختاری شبیه و هماهنگی در مرزهای مشابه رخ می دهد. مایکل اونیل و رایان بر روی عملگرهای ادغام تحقیق کردند و ادغام تک نقطه ای را با دیگر راهکارها و جایگزین ها مقایسه کردند. تحقیق اونیل و رایان پیشنهاد می کند که اساس ادغام تک نقطه ای قابل استفاده در GE سود های زیادی از ادغام مشابه پیچیده را فراهم می کند در حالی که عملکرد سازگارتری از دیگر انواع نشان می دهد. برای این دلایل عملگر ادغام تک نقطه ای برای این تحقیق استفاده می شود.

۳-۱۱-جهش در GE

عملگر جهش در GA ارزش یک کد تنها را در مورد ژنوتیپ دودویی، برای یک ژنوم عدد صحیح شبیه مورد استفاده شده در GE کد داده شده با یک کد دیگر به صورت تصادفی جایگزین می شود. نتیجه جهش در فنوتیپ خیلی عمیق تر و متفاوت تر از ژنو تیپ است. جهش در GE ممکن پتانسیل تخریبی بالایی داشته باشد، مخصوصا زمانی که جهش در حوالی ژنوم رخ دهد.

۳-۱۲-ساختار الگوریتم تکامل دستوری

به طور کلی، الگوریتم تکامل دستوری همانند الگوریتم های ژنتیکی از اجزاء زیر تشکیل می شوند:

¹ Banzhaf
² Conrads
³ Francone
⁴ Nordin
⁵ homologous

۳-۱۲-۱-کروموزوم

در الگوریتم تکامل دستوری، هر کروموزوم نشان دهنده یک نقطه در فضای جستجو و یک راه حل ممکن برای مسئله مورد نظر است. خود کروموزومها (راه حلها) از تعداد ثابتی ژن (متغیر) تشکیل می شوند. برای نمایش کروموزومها، معمولاً از کدگذاریهای دودویی (رشتههای بیتی) استفاده می شود.

۳-۱۲-۲-جمعیت

مجموعه‌ای از کروموزومها یک جمعیت را تشکیل می دهند. با تاثیر عملگرهای دستوری بر روی هر جمعیت، جمعیت جدیدی با همان تعداد کروموزوم تشکیل می شود.

۳-۱۲-۳-تابع برازندگی

به منظور حل هر مسئله با استفاده از الگوریتم تکامل دستوری، ابتدا باید یک تابع برازندگی برای آن مسئله ابداع شود. برای هر کروموزوم، این تابع عددی غیر منفی را برمی گرداند که نشان دهنده شایستگی یا توانایی فردی آن کروموزوم است.

۳-۱۲-۴-عملگرهای الگوریتم تکامل دستوری

در الگوریتم تکامل دستوری، در طی مرحله تولید مثل از عملگرهای دستوری استفاده می شود. با تاثیر این عملگرها بر روی یک جمعیت، نسل بعدی آن جمعیت تولید می شود. عملگرهای انتخاب، آمیزش و جهش معمولاً بیشترین کاربرد را در الگوریتم دستوری دارند.

◆ عملگر انتخاب: (Selection)

این عملگر از بین کروموزوم‌های موجود در یک جمعیت، تعدادی کروموزوم را برای تولید مثل انتخاب می‌کند. کروموزوم‌های برانده‌تر شانس بیشتری دارند تا برای تولید مثل انتخاب شوند.

روش های انتخاب:

◆ انتخاب نخبگان (Elitist Selection):

مناسب‌ترین عضو هر اجتماع انتخاب می‌شود. با توجه به مقدار شایستگی که از تابع ارزیاب دریافت کرده است.

◆ نمونه‌برداری به روش چرخ رولت

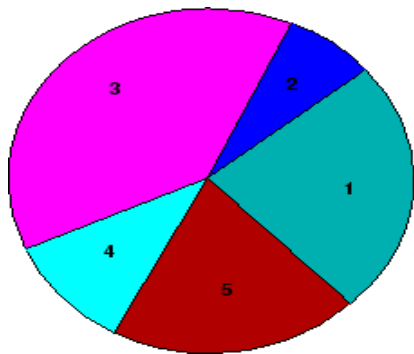
در این روش، به هر فرد قطعه‌ای از یک چرخ رولت مدور اختصاص داده می‌شود. اندازه این قطعه متناسب با برانندگی آن فرد است. چرخ N بار چرخانده می‌شود که N تعداد افراد در جمعیت است. در هر چرخش، فرد زیر نشانگر چرخ انتخاب می‌شود و در مخزن والدین نسل بعد قرار می‌گیرد. این روش می‌تواند به صورت زیر پیاده‌سازی شود:

۱- نرخ انتظار کل افراد جمعیت را جمع کنید و حاصل آن را T بنامید.

۲- مراحل زیر را N بار تکرار کنید:

یک عدد تصادفی r بین 0 و T انتخاب کنید.

در میان افراد جمعیت بگردید و نرخ‌های انتظار (مقدار شایستگی) آنها را با هم جمع کنید تا این که مجموع بزرگتر یا مساوی r شود. فردی که نرخ انتظارش باعث بیشتر شدن جمع از این حد می‌شود، به عنوان فرد برگزیده انتخاب می‌شود.



Population	Fitness
1	25.0
2	5.0
3	40.0
4	10.0
5	20.0

شکل ۳-۷: نحوه ارزیابی شایستگی در چرخ رولت

♦ انتخاب تورنومنت:

یک زیر مجموعه از صفات یک جامعه انتخاب می‌شوند و اعضای آن مجموعه با هم رقابت می‌کنند و سرانجام فقط یک صفت از هر زیرگروه برای تولید انتخاب می‌شوند.

♦ عملگر آمیزش:

در جریان عمل تلفیق به صورت اتفاقی بخشهایی از کروموزوم‌ها با یکدیگر تعویض می‌شوند. این موضوع باعث می‌شود که فرزندان ترکیبی از خصوصیات والدین خود را به همراه داشته باشند و دقیقاً مشابه یکی از والدین نباشند.

هدف تولید فرزند جدید می‌باشد به این امید که خصوصیات خوب دو موجود در فرزندشان جمع شده و یک موجود بهتری را تولید کند.

روش کار به صورت زیر است:

بصورت تصادفی یک نقطه از کروموزوم را انتخاب می‌کنیم

ژن های مابعد آن نقطه از کروموزوم ها را جابجا می کنیم

◆ تلفیق تک نقطه ای :

اگر عملیات تلفیق را در یک نقطه انجام دهیم به آن تلفیق تک نقطه ای می گویند.

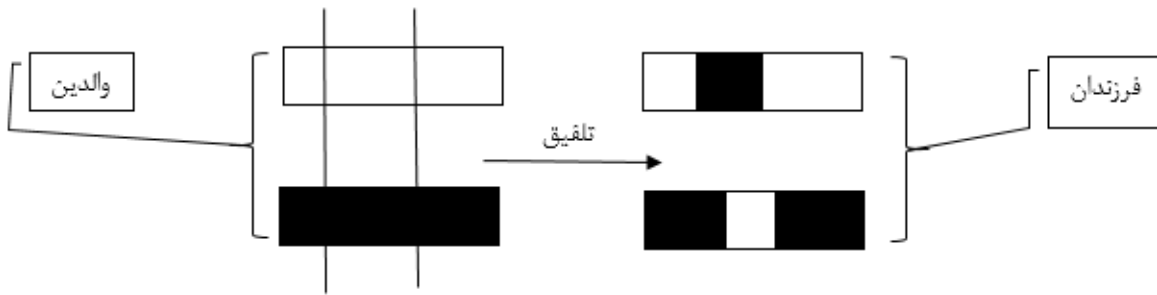
تلفیق بدین صورت انجام می گیرد که حاصل ترکیب کروموزومهای پدر و مادر می باشد. روش تولید مثل نیز بدین صورت است که ابتدا بصورت تصادفی، نقطه ای که قرار است تولید مثل از آنجا آغاز گردد، انتخاب می گردد. سپس اعداد بعد از آن به ترتیب از بیت های کروموزومهای پدر و مادر قرار می گیرد که در شکل زیر نیز نشان داده شده است .

کروموزوم ۱	11011 00100110110
کروموزوم ۲	11011 1100001110
نسل ۱	11011 11000011110
نسل ۲	11011 00100110110

شکل ۳-۸: یک نمونه از تلفیق (آمیزش)

در شکل بالا کروموزوم های ۱ و ۲ در نقش والدین هستند. و حاصل تولید مثل آنها در رشته هائی بنام نسل ذخیره شده است. دقت شود که علامت " | " مربوط به نقطه شروع تولید مثل می باشد و در رشته های نسل اعدادی که بعد از نقطه شروع تولید مثل قرار می گیرند مربوط به کروموزومهای مربوط به خود می باشند. بطوریکه اعداد بعد از نقطه شروع مربوط به نسل ۱ مربوط به اعداد بعد از نقطه شروع مربوط به کروموزوم ۱ و اعداد بعد از نقطه شروع تولیدمثل مربوط به نسل ۲ مربوط به اعداد بعد از نقطه شروع تولید مثل مربوط به کروموزوم ۲ می باشند.

♦ روش ادغام دو نقطه ای:



شکل ۳-۹: ادغام دو نقطه ای

در این روش دو مکان را به صورت تصادفی انتخاب کرده و مقادیر بین این دو نقطه را جابجا می کنیم.

♦ تلفیق نقطه ای:

می توانیم این عملیات را در چند نقطه انجام دهیم ، که به آن بازترکیبی چند نقطه ای می گویند

♦ تلفیق جامع:

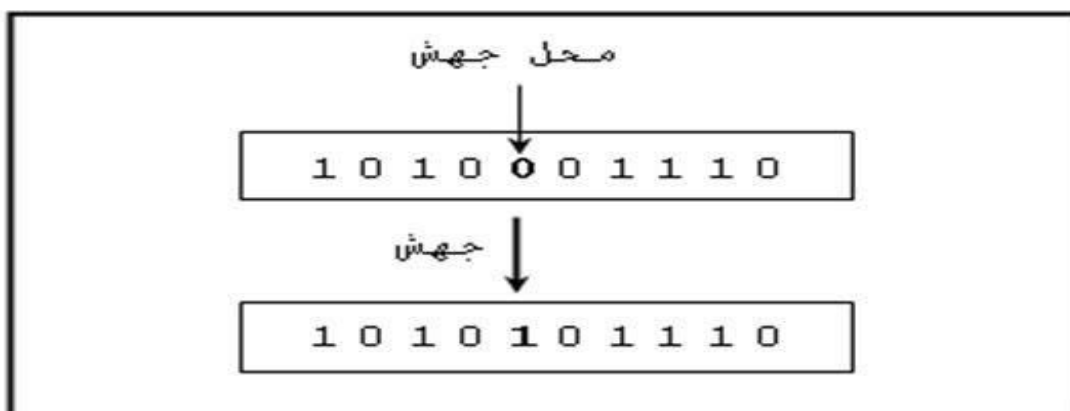
اگر تمام نقاط کروموزوم را بعنوان نقاط بازترکیبی انتخاب کنیم به آن بازترکیبی جامع می گوئیم. مثال آن در شکل ۴ آورده شده است.



شکل ۳-۱۰: تلفیق جامع

◆ عملگر جهش:

پس از اتمام عمل آمیزش، عملگر جهش بر روی کروموزومها اثر داده می‌شود. این عملگر یک ژن از یک کروموزوم را به طور تصادفی انتخاب نموده و سپس محتوای آن ژن را تغییر می‌دهد. اگر ژن از جنس اعداد دودویی باشد، آن را به وارونش تبدیل می‌کند و چنانچه متعلق به یک مجموعه باشد، مقدار یا عنصر دیگری از آن مجموعه را به جای آن ژن قرار می‌دهد. در شکل ۲ چگونگی جهش یافتن پنجمین ژن یک کروموزوم نشان داده شده است.



شکل ۳-۱: یک کروموزوم قبل و بعد از اعمال

پس از اتمام عمل جهش، کروموزومهای تولید شده به عنوان نسل جدید شناخته شده و برای دور بعد اجرای الگوریتم ارسال می‌شوند.

۳-۱۲-۵- مزیت های الگوریتم تکامل دستوری نسبت به الگوریتم ژنتیک

تکامل دستوری براساس اصول سه زمینه مختلف است: برنامه نویسی تکامل خودکار، زیست شناسی مولکولی، و گرامر (اونیل و همکاران ۲۰۰۳). با وجود اینکه تکامل گرامری شکلی از برنامه نویسی ژنتیک است اما با برنامه نویسی ژنتیک سنتی از سه جهت متفاوت است (کوزا، ۱۹۹۲).

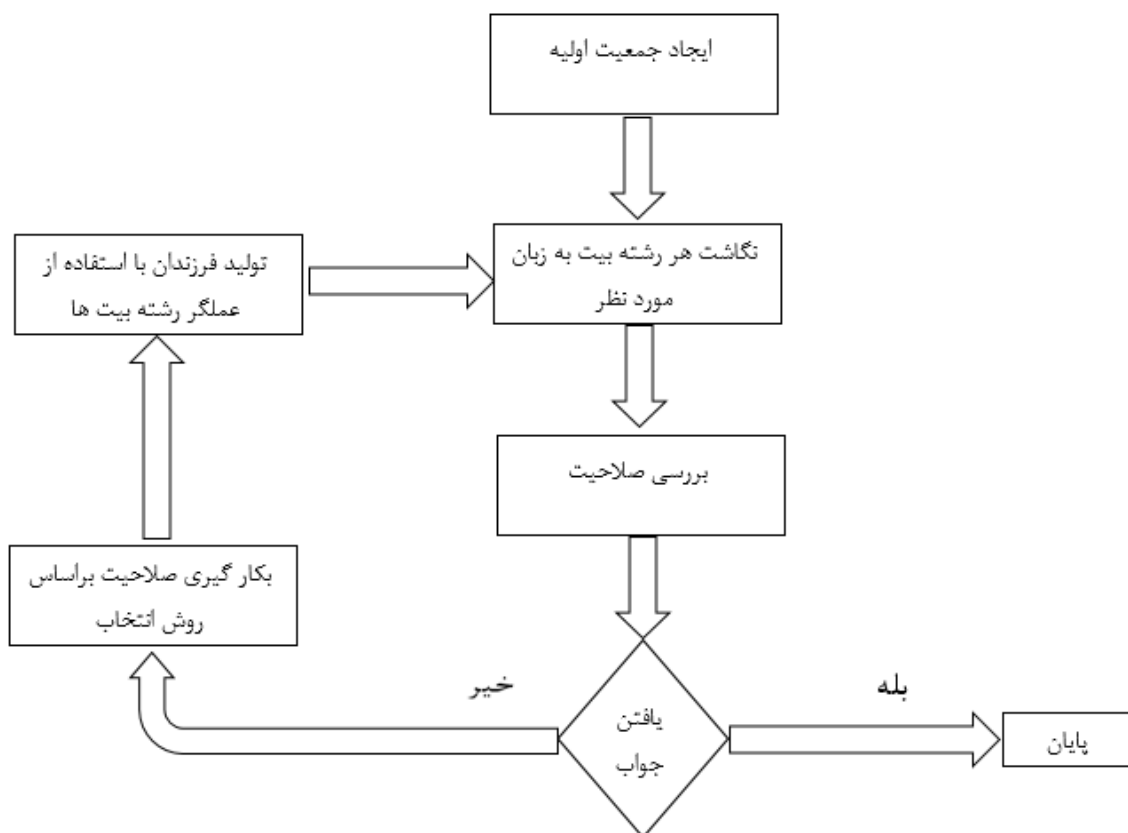
الف) ژنوم های خطی را بکار می گیرد؛ ب) نگاهت وابسته به رشد شناسی از ژنوتیپ به فنوتیپ را اجرا می کند؛ و پ) از یک گرامر به منظور دیکته کردن ساختار های قانونی در فضای فنوتیپی استفاده می کند. تکامل گرامری بیشتر وابسته به فرآیندهای تکاملی مبتنی بر DNA طبیعی هستند تا به برنامه نویسی ژنتیک. زیبایی این الگوریتم در این است که مکانیزم ساده ای ارائه می کند که قابل استفاده برای شرح هر ساختار پیچیده ای مانند زبان ها، نمودارها، شبکه های عصبی، عبارات ریاضی، ترکیبات مولکولی، و غیره است. این ویژگی تکامل گرامری است که آن را به ابزاری قدرتمند برای تکامل هر ساختار دلخواه تا زمانی می کند که این ساختار را بتوان با یک گرامر مستقل از متن تعریف کرد. موارد دیگر که به صورت خلاصه مزیت های الگوریتم تکامل دستوری نسبت به الگوریتم ژنتیک بیان شده است در زیر بیان شده است.

- با پروسه بیوهوژیکی کاملاً تطابق دارد (تطابق کامل با پروسه ژنوتیب - فنو تیپ در سیستم بیولوژیک).
- پیاده سازی الگوریتم تکاملی کاملاً ساده صورت می گیرد (مهم تعریف گرامر و تعریف پروسه نگاهت است).
- کروموزوم های باینری با طول متغیر هستند (در GP درخت داشتیم و درخت ها شکل مختلفی داشتند ولی در اینجا فقط طول کروموزوم ها تغییر پیدا می کنند).
- از فرم BNF استفاده می کند و می تواند به هر زبان دلخواهی برنامه تولید کند.
- از آخرین پیشرفت الگوریتم تکاملی استفاده می کند.
- شیوه تکامل آن از شیوه نگاهت مستقل است.

۳-۱۲-۶- روند کلی الگوریتم تکامل دستوری

قبل از این که یک الگوریتم تکامل دستوری بتواند اجرا شود، ابتدا باید کدگذاری (یا نمایش) مناسبی برای مسئله مورد نظر پیدا شود. معمولی ترین شیوه نمایش کروموزومها در این الگوریتم به شکل

رشته های دودویی است. هر متغیر تصمیم گیری به صورت دودویی در آمده و سپس با کنار هم قرار گرفتن این متغیرها کروموزوم ایجاد می شود. گرچه این روش گسترده ترین شیوه کدگذاری است اما شیوه های دیگری مثل نمایش با اعداد حقیقی در حال گسترش هستند. همچنین یک تابع برازندگی نیز باید ابداع شود تا به هر راه حل کدگذاری شده ارزشی را نسبت دهد. در طی اجراء والدین برای تولید مثل انتخاب می شوند و با استفاده از عملگرهای آمیزش و جهش با هم ترکیب می شوند تا فرزندان جدیدی تولید کنند. این فرآیند چندین بار تکرار می شود تا نسل بعدی جمعیت تولید شود. سپس این جمعیت بررسی می شود و در صورتی که ضوابط همگرایی آورده شوند، فرآیند فوق خاتمه می یابد.



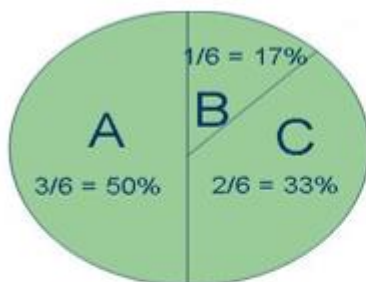
شکل ۳-۱۲: شماتیک ساده و کلی از الگوریتم تکامل دستوری

۳-۱۲-۷- روند کلی بهینه سازی و حل مسائل در الگوریتم تکامل دستوری:

۱- شروع : (Start) تولید تصادفی یک جمعیت (Population) که شامل تعداد زیادی

کروموزم (روشهای حل مسئله است) می باشد.

۲- صحت و درستی : (Fitness) ارزیابی صحت برای تابع $f(x)$ به ازای هر کروموزوم x در جمعیت.



$$\text{fitness}(A) = 3$$

$$\text{fitness}(B) = 1$$

$$\text{fitness}(C) = 2$$

شکل ۳-۱۳: نحوه ارزیابی تابع شایستگی در چرخ رولت (براساس الگوریتم ژنتیک)

۳- ایجاد یک جمعیت جدید: (New Population) تولید یک جمعیت جدید با انجام تمامی زیر

گروههای زیر تا آنکه یک جمعیت جدید ایجاد گردد.

۱-۳- انتخاب: (Selection) انتخاب کروموزومهای پدر و مادر از جمعیت قبلی با توجه به صحت و درستی آن. (Fitness) بطوریکه هر چه Fitness بهتر باشد (دقت جواب در همگرایی بیشتر باشد) شانس بیشتری برای انتخاب دارد.

۲-۳- تولید مثل: (Crossover) انجام زادو ولد و ایجاد یک نسل جدید.

۳-۳- مشخص شدن مکان فرزند تولید شده در کروموزوم

۴-۳- پذیرش: (Accepting) جا دادن فرزند جدید در داخل جمعیت.

۵-۳- جایگزینی: (Replace) جایگزینی جمعیت جدید به جای جمعیت قبلی و مورد استفاده قرار دادن جمعیت جدید در مراحل بعدی الگوریتم

۵-۳- امتحان: (Test): اگر شرایط مطلوب در حل مسئله ارضا شد اعلام میکنیم که به بهترین جواب رسیده ایم و از الگوریتم خارج می شویم در غیر این صورت به مرحله ۲ یعنی Fitness میرویم و دوباره همین روند را تکرار می کنیم.

۳-۱۲-۸- شرط پایان الگوریتم

چون که الگوریتم های تکامل دستوری بر پایه تولید و تست می باشند، جواب مساله مشخص نیست و نمی دانیم که کدامیک از جواب های تولید شده جواب مطلوب است تا شرط خاتمه را پیدا شدن جواب در جمعیت تعریف کنیم. به همین دلیل، معیارهای دیگری را برای شرط خاتمه در نظر می گیریم:

۱. تعداد مشخصی نسل: می توانیم شرط خاتمه را مثلاً ۱۰۰ دور چرخش حلقه اصلی برنامه قرار دهیم .

۲. عدم بهبود در بهترین شایستگی جمعیت در طی چند نسل متوالی

۳. بهترین شایستگی جمعیت تا یک زمان خاصی تغییری نکند.

شرایط دیگری نیز می‌توانیم تعریف کنیم و همچنین می‌توانیم ترکیبی از موارد فوق را به عنوان شرط خاتمه به کار ببندیم .

نسل بعدی: دو ژنوم (کروموزوم) به طور تصادفی برای تولید نسل بعدی انتخاب می‌شوند.

۱-انتخاب

دو ژنوم به طور تصادفی از نسل قبل انتخاب می‌شوند. این ژنوم‌ها دارای اعداد تناسب بزرگتری هستند و بعضی صفات آنها به نسل بعدی منتقل می‌شوند. این بدین معنی است که عدد تناسب در حال افزایش خواهد بود.

۲-تلفیق (Crossover)

حالا دو ژنوم بخشی از ژنهایشان را برای ایجاد نسل بعدی اهدا می‌کنند. اگر آنها تغییر پیدا نکنند همانطور بی‌تغییر به نسل بعدی منتقل خواهند شد. درجه Crossover نشان دهنده این است که هر چند وقت یکبار ژنوم‌ها تغییر پیدا خواهند کرد.

عملگر تغییر در ژنوم‌های باینری مثال ما با انتخاب یک مکان تصادفی در ژنوم برای تغییر آغاز می‌شود. بخش اول ژنهای پدر و بخش دوم ژنهای مادر با هم ترکیب می‌شوند (و بالعکس) تا ۲ فرزند تولید شوند. در زیریک عمل تغییر را می‌بینیم.

قبل از تلفیق

پدر ۰۱۱۱۱۰۰۱۰۰۱۱۰۰۱۰۱۱۰۱۱۰۰۰۱۱۱۰۱۱۰۱۰۰۰

مادر	۰۱۰۱۰۰۱۰۰۱۱۰۰۱۰۱۱۰۱۱۰۰۰۱۱۱۰۱۱۰۱۰۰۰۰
------	-------------------------------------

جدول ۳-۱: قبل از تلفیق

بعد از تلفیق

فرزند ۱ ۰۱۱۱۱۰۰۱۰۰۱۱۰۱۰۱۱۱۱۱۱۰۱۰۰۰۱۰۰۰۱۰۰۱۰

فرزند ۲	۰۱۰۱۰۰۱۱۱۱۱۰۰۰۱۰۱۱۰۱۱۰۰۰۱۱۱۰۱۱۰۱۰۰۰۰
---------	--------------------------------------

جدول ۳-۲: عمل بعد از تلفیق

۳- جهش (Mutation)

قبل از این که ژنوم ها در نسل بعدی قرار بگیرند، احتمال دارد دچار جهش یا تغییر ناگهانی شوند. جهش یک تغییر ناگهانی در ژن است. در ژنهای باینری این تغییر به معنای تغییر یک بیت از ۰ به ۱ یا از ۱ به ۰ است. درجه جهش نشان دهنده احتمال بروز جهش در یک ژن است. این روند تا تولید نسل های متعددی ادامه می یابد تا در نهایت به جواب برسیم.

۳-۱۳- نتیجه گیری

الگوریتم تکامل دستوری از جمله الگوریتم‌هایی است که دارای قدرت بسیار زیادی در یافتن جواب مسئله هستند، اما باید توجه داشت که شاید بتوان کاربرد اصلی این الگوریتم ها را در مسائلی در نظر گرفت که دارای فضای حالت بسیار بزرگ هستند و عملاً بررسی همه حالت‌ها برای انسان در زمان‌های نرمال (در حد عمر بشر) ممکن نیست. از طرفی باید توجه داشت که حتماً بین حالات مختلف مسئله

باید دارای پیوستگی مناسب و منطقی باشیم. در نهایت الگوریتم تکامل دستوری این امکان را به ما می‌دهد که دارای حرکتی سریع در فضای مسئله به سوی هدف باشیم. به گونه‌ای که می‌توانیم تصور کنیم که در فضای حالات مسئله به سوی جواب مشغول پرواز هستیم.

۶- فصل چہارم

نتیجہ پژوهش

۴-۱-مقدمه

این نوع پژوهش از نوع پژوهش های کمی بوده و با توجه به بکارگیری الگوریتم تکامل دستوری در پیش‌بینی زمینه ورشکستگی و نتایج مطلوب و مناسب، در اینجا از الگوریتم تکامل دستوری و مراحل اجرایی آن به صورت تفصیلی آورده شده است. در این پژوهش از ۲۲ نوع شاخص به عنوان متغیرهای مستقل جهت پیش‌بینی شاخص کل استفاده شده است. مراحل اجرای فرایند بکار گیری داده ها و تحلیل آنها در ادامه آورده شده است.

۴-۲-تعریف مسئله

تعریف مساله در فصل معرفی تحقیق تشریح گردیده است و مشخص گردیده است که هدف تحقیق چه می باشد.

۴-۳-انتخاب متغیرها

انتخاب متغیرها از گروه های مشخص شده در سازمان بورس و شاخص های معرفی شده در بازه زمانی صورت گرفته است. در این تحقیق ۲۲ شاخص در نظر گرفته بعنوان ۲۲ متغیر مستقل بکار گرفته شده است و به همراه نماد جهت استفاده در کدنویسی و مطالب بعد در اینجا آورده شده است. قابل ذکر است که شماره های در ردیف آمده بعنوان شماره متغیر در ادامه استفاده می‌شود.

ردیف	شاخص	نماد	ردیف	شاخص	نماد
۱	۵۰ شرکت برتر	Var1	۱۲	ماشین الات و تجهیزات	Var12
۲	کل	Var2	۱۳	سرمایه گذاری ها	Var13
۳	بازار دوم	Var3	۱۴	فرآورده های نفتی و کک	Var14
۴	۳۰ شرکت بزرگ	Var4	۱۵	بانک ها و موسسات اعتباری	Var15
۵	آزاد شناور	Var5	۱۶	بازار اول	Var16

Var17	قیمت و بازده نقدی	۱۷	Var6	خودرو و ساخت قطعات	۶
Var18	واسطه گری پولی و مالی	۱۸	Var7	زراعت و خدمات وابسته	۷
Var19	فلزات اساسی	۱۹	Var8	محصولات شیمیایی	۸
Var20	حمل و نقل، انبار داری و ارتباطات	۲۰	Var9	محصولات چوبی	۹
Var21	لاستیک و پلاستیک	۲۱	Var10	صنعت	۱۰
Var22	مواد و محصولات دارویی	۲۲	Var11	ساخت محصولات فلزی	۱۱

در ابتدا تغییرات ۲۰۰ روز شاخص های ذکر شده (۱۰۰ روز کاهش شاخص نسبت به روز قبل و ۱۰۰ روز افزایش نسبت به روز قبل) طبق آخرین تغییرات در نظر گرفته شده است. اما در ادامه این تعداد به ۸۰ روز افزایشی و ۶۰ روز کاهشی، جهت افزایش دقت کاهش پیدا کرد همچنین در ادامه یا با بکار گیری نام لاتین و یا نمادها، در این پژوهش استفاده می شود.

۴-۴- تحلیل داده ها

در این مرحله داده های شاخص کل بورس اوراق بهادار تهران از سال از ابتدای سال ۱۳۸۷ تا انتهای مرداد ماه سال ۱۳۹۵ و دیگر شاخص ها نیز طبق آخرین تغییرات مورد نیاز گردآوری شده است. داده های مذکور و منابع مورد نیاز مطابق تأیید و اعتباردهی سازمان بورس و اوراق بهادار تهران است.

۴-۵- آماده سازی داده ها

در این تحقیق در این مرحله داده های گردآوری شده در مرحله قبل مورد بررسی توصیفی قرار خواهد گرفت. ماهیت داده ها، گرانولیتته داده ها و کیفیت داده ها موارد مورد بررسی در این مرحله می باشند. همچنین در این مرحله کاهش داده ها^۱ صورت خواهد پذیرفت. بدین منظور با توجه به دامنه تحقیق،

^۱ Data Reduction

مطالعه و کنکاشی در روشهای شناسایی ویژگی و یا استخراج ویژگی صورت نخواهد گرفت ولی با ارجاع به تحقیقات پیشین از نتایج صحنه گذاری شده قبلی بهره گرفته خواهد شد.

۴-۶-مدل سازی

در این مرحله تقسیم بندی داده ها به مجموعه داده های آموزش و مجموعه داده های تست انجام می شود. بدیهی است با توجه به آنکه حجم داده های مورد بررسی بسیار بزرگ است تخصیص قسمتی از داده ها به داده های صحنه گذاری ضروری می باشد. مدل به کارگرفته شده جهت پیش بینی، مدل الگوریتم تکامل دستوری و شبکه عصبی مصنوعی پیش رو می باشد که آموزش آن به کمک الگوریتم های تکاملی صورت خواهد پذیرفت. از الگوریتمهای ژنتیک به منظور آموزش شبکه عصبی مصنوعی به کار گرفته خواهند شد. بدین منظور از نرم افزار MATLAB جهت برنامه نویسی استفاده می شود.

۴-۶-۱-کد نویسی

در این مرحله اطلاعات لازم برای ساخت مدل به برنامه متلب فراخوانی شد و روند زیر طی گردید. شاخص های مربوط به دو دسته تقسیم شده و هر شاخص کاهشی معادل یک شاخص افزایشی در نظر گرفته می شود. ۷۵ درصد داده ها یعنی ۷۵ عدد از شاخص ها (کاهشی و افزایشی) برای آموزش شبکه بکار گرفته شده اند و ۲۵ درصد یعنی ۲۵ شاخص برای آزمودن شبکه آموزش داده شده استفاده شده است. ضرایب اولیه در بازه ۱۰۰۰۰- تا ۱۰۰۰۰ انتخاب شده اند، به گونه ای که چند جمله ای ایجاد می شود که از ترکیب اعداد ۱۰۰۰۰-، ۱۰۰۰-، ۱۰۰-، ۱۰-، ۱، ۰، ۱، ۱۰، ۱۰۰، ۱۰۰۰، ۱۰۰۰۰ و ضرب این اعداد رندوم به دست می آیند.

```
Randompool = [-10000*rand,-1000*rand,-100*rand,-
```

```
10*rand, 0, 10 *rand, 100*rand, 1000*rand, 10000*rand];
```

این ضرایب باید در تک تک شاخص ها ضرب شوند تا خروجی به شکل زیر در آید.

$$\text{Output} = (\langle \text{some expression} \rangle * \text{var1}) + \dots + (\langle \text{Some expression} \rangle * \text{var22})$$

در دور اول اجرا به تعداد ۱۰۰۰ iteration^۱ اجرا شده اند تا نقطه تمایز دسته بندی مناسب (Differentiation Point) مشخص شود. نقاط تمایز هم در بازه ۵۰۰۰- تا ۵۰۰۰ بررسی شده اند. با تخصیص نقطه تمایز از دور اول برای هر یک از سال ها با ۷۵ درصد از داده ها ۲۰ بار با تعداد iteration ۲۰۰۰ اجرا گردیده اند، و میانگین ضرایب به دست آمده اند. در این مرحله تعدادی از شاخص های بی اثر حذف می شوند. پس از انتخاب نقطه جدایی مناسب، برنامه با نقاط جدایی تعیین شده برای هر کدام از دوره های مدنظر که در اینجا ۴ بار اجرا می شوند و ضرایب نهایی با هم مقایسه شده اند. هر یک از ضرایب در هر مرحله از اجرا حداقل یک بار صفر شده باشد حذف می شود. جدول زیر هر یک از شاخص هایی که حین اجرای برنامه برای پیش بینی روز های آتی حذف شده اند، مشخص شده است.

جدول ۴-۱: متغیرهای حذف شده حین اجرای برنامه

شاخص	t+1	t+2	t+3
Var1			
Var2			
Var3			
Var4			
Var5			
Var6			
Var7			

^۱ تکرار

Var8			
Var9			
Var10			
Var11			
Var12			
Var13			
Var14			
Var15			
Var16			
Var17			
Var18			
Var19			
Var20			
Var21			
Var22			

جهت پیش‌بینی یک روز بعد ($t+1$) متغیر های ۲، ۱۰، ۱۳، ۱۷ و ۱۸ بعنوان متغیر های اثرگذار و مفید در نظر گرفته می‌شود. برای پیش‌بینی دو روز بعد متغیر های ۳، ۱۰، ۱۶، ۱۷ و ۱۸، برای پیش‌بینی سه روز بعد متغیر های ۱، ۴، ۱۴، ۱۵ و ۱۸ اثرگذار هستند.

برای هر یک روزی که باید پیش‌بینی شود، برنامه سه بار اجرا می‌شود و نسبت های با ضریب صفر یا خیلی کم حذف می‌شوند و برنامه با تعداد دور بالا اجرا می‌شود تا دقت بهتری به دست آید. در نهایت بهترین نقاط تمایز و مناسب ترین شاخص ها و بهترین ضرایب گزارش می گردد. در دور سوم برای هر یک

از روزها برنامه ای جدا نوشته می‌شود و شاخص های با ضریف صفر حذف شده اند تا دقت بالاتری به دست آید.

۴-۷-نتایج به دست آمده از الگوریتم تکامل دستوری در نرم افزار متلب

۴-۷-۱-نتایج حاصل برای یک روز آتی

جدول ۴-۲:نتایج پیش‌بینی برای یک روز آینده همراه با متغیرهای مؤثر

Diff point	100	
Iterations	10000	
Test result	32	82%
Var2	-1139.0925	
Var10	-116.0724	
Var13	984.2418	
Var17	39.0907	
Var18	83.3865	
Fitness	109	87%
Overall	140	89%

۴-۷-۲-نتایج حاصل برای روز دوم

جدول ۴-۳:نتایج پیش‌بینی برای روز دوم همراه با متغیرهای مؤثر

Diff point	300	
Iterations	10000	
Test result	34	85%
Var3	-7.1122	
Var10	1.0482	
Var16	135.4782	
Var17	59.1584	

Var18	-26.7194	
Fitness	120	82%
Overall	138	86%

۴-۷-۳- نتایج حاصل برای روز سوم

جدول ۴-۴: نتایج پیش‌بینی برای روز سوم همراه با متغیرهای مؤثر

Diff point	300	
Iteration	10000	
Test result	33	81%
Var1	22.2401	
Var4	-86.0334	
Var14	98.3186	
Var15	112.2599	
Var18	46.6862	
Fitness	104	86%
Overall	130	87%

۴-۸- نقطه تمایز افزایش یا کاهش شاخص در جدول زیر ارائه می‌شود.

جدول ۴-۵: شرایط افزایش یا کاهش شاخص

روز	نقطه تمایز افزایش یا کاهش شاخص
t+1	شاخص کل برای یک روز آتی افزایش می‌یابد ← اگر نقطه تمایز < 100 شاخص کل برای یک روز آتی کاهش می‌یابد ← اگر نقطه تمایز > 100
t+2	شاخص کل در روز دوم آتی افزایش می‌یابد ← اگر نقطه تمایز < 300 شاخص کل در روز دوم آتی کاهش می‌یابد ← اگر نقطه تمایز > 300
t+3	شاخص کل در روز سوم آتی افزایش می‌یابد ← اگر نقطه تمایز < 300 شاخص کل در روز سوم آتی کاهش می‌یابد ← اگر نقطه تمایز > 300

۴-۹- شاخص های اثر گذار برای هر روز در جدول زیر آمده است.

جدول ۴-۶: شاخص های اثر گذار برای پیش بینی هر روز

روز	شاخص های اثر گذار
t+1	$-1139.0925\text{var}^*2 - 116.0724\text{var}^*10 + 984.2418\text{var}^*13 + 39.0607\text{var}^*17 + 83.3865\text{var}^*1$
t+2	$-7.1122\text{var}^*3 + 1.0482\text{var}^*10 + 135.4782\text{var}^*16 + 59.1584\text{var}^*17 - 26.7197\text{var}^*18$
t+3	$22.2401\text{var}^*1 + -86.0334\text{var}^*4 + 98.3186\text{var}^*14 + 112.2599\text{var}^*15 + 46.6862\text{var}^*18$

۴-۹-۱- مدل ارائه شده برای پیش بینی یک روز آتی

$$F_1 = -1139.0925\text{var}^*2 - 116.0724\text{var}^*10 + 984.2418\text{var}^*13 + 39.0607\text{var}^*17 + 83.3865\text{var}^*18$$

طبق معادله بالا اگر F_1 بزرگتر از ۱۰۰ باشد پیش بینی می شود طی یک روز آینده شاخص افزایش می یابد و در صورتی که کوچکتر از ۱۰۰ باشد با کاهش شاخص بازار مواجه می شود.

۴-۹-۲- مدل ارائه شده برای پیش بینی روز دوم آتی

$$F_2 = -7.1122\text{var}^*3 + 1.0482\text{var}^*10 + 135.4782\text{var}^*16 + 59.1584\text{var}^*17 - 26.7197\text{var}^*18$$

مدل ارائه شده پیش بینی می کند که در صورتی که F_2 بزرگتر از ۳۰۰ باشد بازار با افزایش شاخص کل و در صورتی که کمتر از ۳۰۰ باشد با کاهش شاخص همراه است.

۴-۹-۳- مدل ارائه شده برای پیش بینی روز سوم آتی

$$F_3 = 22.2401\text{var}^*1 - 86.0334\text{var}^*4 + 98.3186\text{var}^*14 + 112.2599\text{var}^*15 + 46.6862\text{var}^*18$$

*Var: در اینجا var برابر هر یک از شاخص های روزانه تقسیم بر میانگین دوره در نظر گرفته شده است.

طبق معادله ارائه شده ی فوق پیش‌بینی می‌شود که اگر F_3 بزرگتر از ۳۰۰ باشد شاخص کل در روز سوم با افزایش و در صورتی که کمتر از ۳۰۰ باشد با کاهش همراه است.

۴-۱۰-اعتبار سنجی

برای اعتبار سنجی فرمول های به دست آمده، داده های واقعی برای ده روز بعد از داده هایی که به عنوان ورودی در نظر گرفته شده است با فرمول ها مقایسه می‌شود. برای به دست آوردن var^* ، ابتدا میانگین دوره را برای هر یک از شاخص ها به دست می آوریم سپس نسبت هر یک از آنها بر میانگین به دست آمده محاسبه می‌شود. جدول زیر میانگین دوره هر یک از یازده شاخص موثر است.

جدول ۴-۷: میانگین دوره در نظر گرفته شده برای شاخص های مؤثر

شاخص مؤثر	مقدار میانگین دوره
۵۰ شرکت برتر	۲۱۶۸/۱۵۹
کل	۵۲۷۶۲/۹۳۲۵
بازار دوم	۹۶۷۵۱/۶۳۱۵
۳۰ شرکت بزرگ	۲۵۰۴/۵۸۹
صنعت	۴۳۹۶۸/۷۴۴
سرمایه گذاری ها	۹۸۲/۳۸۵۳۵
فرآورده های نفتی و کک	۱۳۹۶۲۷/۲۸۹۳
بانک ها و مؤسسات اعتباری	۶۲۲/۳۲۱
بازار اول	۳۸۵۴۵/۵۳۷
قیمت و بازده نقدی	۲۰۰۸۵۲/۷۴۸
واسطه گری پولی و مالی	۱۰۸۳۸۷/۷۳۵

برای پیش‌بینی روز اول (هفتم دی ماه سال نود و پنج) شاخص های کل، صنعت، سرمایه گذاری ها، قیمت و بازده نقدی و واسطه گری های پولی و مالی به عنوان ورودی فرمول در نظر گرفته می‌شود. برای این کار ابتدا باید نسبت شاخص ها به میانگین دوره خود محاسبه شود که در جدول زیر به دست آمده است.

جدول ۴-۸: مقادیر محاسبه شاخص برای روز اول

شاخص کل	شاخص صنعت	شاخص سرمایه گذاری ها	شاخص قیمت و بازده نقدی	شاخص واسطه گری پولی و مالی
۱/۵۱۵۴۳۷	۱/۵۳۱۵۳۸	۱/۸۵۲۲۳	۱/۵۹۰۴۴۵	۱/۳۹۰۲۶۸

با قرار دادن مقادیر جدول فوق در فرمول F_1 مقدار $۹۷/۱۰۳$ به دست می آید که بعلت کمتر از ۱۰۰ بودن مقدار محاسبه شده نشان دهنده کاهش شاخص در روز آینده (هفتم دی ماه) است.

برای پیش‌بینی روز دوم مقادیر شاخص های بازار دوم، صنعت، بازار اول، قیمت و بازده نقدی و واسطه گری های پولی و مالی بعنوان ورودی فرمول در نظر گرفته می‌شود. ابتدا باید نسبت مقادیر متغیرها به میانگین دوره محاسبه شود.

جدول ۴-۹: مقادیر محاسبه شاخص برای روز دوم

شاخص بازار دوم	شاخص صنعت	شاخص بازار اول	شاخص قیمت و بازده نقدی	شاخص واسطه گری های پولی و مالی
۱/۷۵۰۸۸۵	۱/۵۳۱۵۴۸	۱/۴۷۶۱۷۹	۱/۵۹۰۴۴۵	۱/۳۶۰۲۶۸

با قرار دادن مقادیر فوق در فرمول F_2 مقدار $۲۴۶/۸۸۵۴$ به دست می آید که این امر به دلیل اینکه عدد به دست آمده کمتر از ۳۰۰ است، نشان دهنده کاهش شاخص است.

برای پیش‌بینی روز سوم مقادیر شاخص های ۵۰ شرکت برتر، ۳۰ شرکت بزرگ، فرآورده های نفتی و کک، بان ها و مؤسسات اعتباری و اسطه گری های پولی و مالی به عنوان ورودی فرمول F_3 در نظر گرفته شود. ابتدا مقادیر نسبت شاخص های مؤثر نسبت به میانگین دوره خود در جدول زیر محاسبه شده است.

جدول ۱۰-۴: مقادیر محاسبه شاخص برای روز سوم

۵۰ شرکت برتر	۳۰ شرکت بزرگ	فرآورده های نفتی و کک	بانک ها و مؤسسات اعتباری	واسطه گرهای پولی و مالی
۱/۴۵۳۶۷۶	۱/۳۰۸۵۱۸	۱/۴۳۶۵۸۲	۱/۱۲۶۷۵	۱/۳۹۰۲۶۸

با قرار دادن مقادیر فوق در فرمول F_3 مقدار ۲۵۲/۳۶۳۸ به دست می آید که این مقدار با توجه به اینکه کمتر از ۳۰۰ است، می توان نتیجه گرفت که در روز سوم نیز شاخص کاهش یافته است.

با توجه به مقادیر به دست آمده در روزهای هفتم، هشتم و یازدهم دی ماه نود و پنج نسبت به روز ششم دی ماه با کاهش شاخص همراه بوده است که در مقایسه با واقعیت، با صحت پیش‌بینی همراه است.

برای ۹ روز آتی (برای هر یک از تاریخ های ذکر شده در جدول) با استفاده از الگوریتم تکامل دستوری و فرمولی که به دست آمده و مقایسه آن با مقدار واقعی به طور خلاصه در جدول زیر آورده شده است.

جدول ۱۱-۴: اعتبار سنجی پیش‌بینی با واقعیت (دی ماه ۹۵)

ردیف	روز اول آتی	روز دوم آتی	روز سوم آتی
۱	کاهش ۷۴/۴۰۹	کاهش ۱۸۳/۶۴۸	کاهش ۲۵۱/۶۶
تاریخ	۹۵/۱۰/۰۸	۹۵/۱۰/۱۱	۹۵/۱۰/۱۲
۲	کاهش ۶۴/۷۴	کاهش ۲۴۵/۳۱۹	افزایش ۳۵۱/۷۰۹
تاریخ	۹۵/۱۰/۱۲	۹۵/۱۰/۱۳	۹۵/۱۰/۱۴
۳	کاهش ۶۹/۹۷	افزایش ۳۴۰/۴۹۵	افزایش ۳۶۶/۳۶۳

۹۵/۱۰/۱۵		۹۵/۱۰/۱۴		۹۵/۱۰/۱۳		تاریخ
کاهش	۲۹۱/۵۶۲۳	افزایش	۳۱۸/۴۱۹	افزایش	۱۱۱/۱۹۵	۴
۹۵/۱۰/۱۸		۹۵/۱۰/۱۵		۹۵/۱۰/۱۴		تاریخ
کاهش	۲۶۵/۹۸	کاهش	۲۴۲/۳۲۸	کاهش	۷۸/۷۶	۵
۹۵/۱۰/۱۹		۹۵/۱۰/۱۸		۹۵/۱۰/۱۵		تاریخ
کاهش	۱۹۶/۲۸	کاهش	۲۴۲/۷۶	کاهش	۸۸/۸۴	۶
۹۵/۱۰/۲۰		۹۵/۱۰/۱۹		۹۵/۱۰/۱۸		تاریخ
کاهش	۲۴۴/۵۰۴	کاهش	۲۹۴/۴۶۰۵	افزایش	۱۰۳/۹۴	۷
۹۵/۱۰/۲۱		۹۵/۱۰/۲۰		۹۵/۱۰/۱۹		تاریخ
افزایش	۳۰۹/۴۳۹	کاهش	۲۵۶/۷۸۳	کاهش	۸۹/۳۰۴	۸
۹۵/۱۰/۲۲		۹۵/۱۰/۲۱		۹۵/۱۰/۲۰		تاریخ
افزایش	۳۵۳/۴۰۹	افزایش	۳۱۵/۴۵۳	کاهش	۸۱/۶۷۳	۹
۹۵/۱۰/۲۵		۹۵/۱۰/۲۲		۹۵/۱۰/۲۱		تاریخ

در جدول فوق مواردی که با رنگ سیاه مشخص شده اند نشان دهنده خطای پیش‌بینی است. با توجه به دو حالت (کاهش یا افزایش) شاخص و نمونه ای که جهت مقایسه و اعتبار سنجی با واقعیت در نظر گرفته شده است میزان خطا را نمی‌توان از روش های دیگر به صورت کمی به دست آورد. اما تعداد محاسباتی که جهت ابعاد سنجی صورت گرفته است، ۳۰ مورد است که ۴ مورد آن با خطا مواجه شده است. بنابراین می‌توان میزان خطا را ۱۳/۳۳ درصد در نظر گرفت.

۵- فصل پنجم:

نیچرگیری، پیشنهادات و

محدودیت‌های تحقیق

۵-۱- نتایج حاصل از مدل GE در پیش‌بینی شاخص کل

نتایج به دست آمده از مدل GE، امیدوار کننده است، با وجود انتخاب شاخص های صنایع گوناگون در بورس اوراق بهادار تهران، این مدل بیانگر یک دقت بالای طبقه بندی در نمونه انتخاب شده است. که به آرامی این میزان کم می‌شود، تا اینکه به طور ناگهانی در روزهایی که پیش‌بینی انجام می‌شود کاهش پیدا می‌کند. این الگوریتم این توانایی را دارد که از میان ۲۲ متغیر بهترین متغیر ها را انتخاب می‌کند، اما در می‌یابیم که چون بسیاری از شاخص ها دارای محتوای یکسانی هستند لذا طبقه بندی از طریق ساخت مدل هایی با تعداد زیاد این شاخص ها، ارتقا پیدا نمی‌کند. عوامل ریسک پیشنهادی توسط هر مدل تا حدودی متفاوت می‌باشد و شامل برخی یافته های کمتر بصری اما با این وجود محتمل می‌باشد. با اعتبارسنجی صورت گرفته در فصل چهار می‌توان چنین نتیجه‌گیری کرد که پیش‌بینی با الگوریتم تکامل دستوری از اطمینان و اعتبار بالایی برخوردار است.

۵-۲- بحث و نتیجه گیری

در زمینه پیش‌بینی شاخص بورس الگوریتم تکاملی تا کنون بکار نرفته است. عوامل مؤثر در پیش‌بینی شاخص بورس برای یک روز بعد شاخص های کل، صنعت، سرمایه گذاری ها، قیمت و بازده نقدی و واسطه گری پولی و مالی، برای دو روز بعد شاخص های بازار دوم، صنعت، بازار اول، قیمت و بازده نقدی و واسطه گری پولی و مالی برای سه روز بعد شاخص های ۵۰ شرکت برتر، ۳۰ شرکت بزرگ، فرآورده های نفتی و کک، بانک مؤسسات اعتباری و واسطه گری پولی و مالی و هستند. نتایج این تحقیق به صورت خلاصه در جدول زیر ارائه شده است.

جدول ۵-۱: متغیرهای مؤثر و ریسک‌زا برای هر یک از روزهای پیش‌بینی شده

روزهای پیش‌بینی شاخص	عوامل ریسک‌زا
یک روز بعد	شاخص کل - شاخص صنعت - شاخص سرمایه‌گذاری‌ها - شاخص قیمت و بازده نقدی - واسطه‌گری پولی و مالی
دو روز بعد	شاخص بازار دوم - شاخص صنعت - شاخص بازار اول - شاخص قیمت و بازده نقدی - شاخص واسطه‌گری پولی و مالی
سه روز بعد	شاخص پنجاه شرکت برتر - شاخص سی شرکت برتر - شاخص نفتی و و فرآورده‌های کک - شاخص بانک و مؤسسات اعتباری - شاخص واسطه‌گری پولی و مالی

۵-۳- محدودیت‌ها

در این پژوهش با محدودیت‌های زیر همراه بوده است.

- روزهایی که شاخص با افت شدید یا کاهش شدید همراه است به عنوان ورودی در نظر گرفته نشده است.
- شاخص‌روزی که بنا به اتفاقات سیاسی یا اتفاقات خارج حیطه بورس و بازار مالی، در صورتی که همراه با تغییرات زیاد بوده است در اطلاعات ورودی گنجانده نشده است.
- توان محاسباتی نرم‌افزار مورد استفاده که در این تحقیق نرم‌افزار MATLAB می‌باشد یکی دیگر از موارد مورد توجه می‌باشد. این نرم‌افزار بسیار قدرتمند در انجام محاسبات، توان محدودی نسبت به برخی نرم‌افزارهای موجود مانند R در راندمان محاسبات و فضای مورد نیاز دارد. این مورد اثری در محاسبات و نتایج این تحقیق ایجاد ننموده لکن امکان محاسبات همزمان تعداد نمونه بیشتر را با محدودیت مواجه می‌نماید.

۵-۴-پیشنهادهات

- ۱- در زمینه های کاربردی فراوانی می توان برای پیش بینی از الگوریتم تکامل دستوری از جمله سیستم بازرگانی، مدل سازی اقتصادی، خرید و فروش ارز، تضمین وام و ... استفاده کرد. به سرمایه گذاران، بانک ها، کسانی که در بازار بورس فعالیت دارند، تصمیم گیرندگان خرید سهام شرکت ها، اعطا کنندگان وام به شرکت، همنین جهت ارزیابی عملکرد و پیش بینی دوام شاخص توصیه می شود که از این تحقیق استفاده کنند.
- ۲- می توان مدل تکامل دستوری بدست آمده از این تحقیق بر روی نمونه های دیگر بررسی و نتایج حاصل از آن با دیگر روش های پیش بینی شاخص، همچون شبکه عصبی، الگوریتم ژنتیک و مدل های سابق مقایسه کرد.
- ۳- مدل الگوریتم تکاملی را می توان در زمینه های دیگر پیش بینی مانند آب و هوا و ... به طور جداگانه بکار بست و با دیگر روش ها مقایسه نمود.
- ۴- هنگام پیش بینی روزها براساس روزی که پیش بینی می کنیم افزایش یا کاهش سه روز آینده را نظر می گیریم (بطور مثال ممکن است روز سوم نسبت به روز دوم با افزایش شاخص همراه باشد).

منابع و مأخذ

۱. آذر، ع. و افسر، ا. (۱۳۸۵). مقایسه روش های کلاسیک و هوش مصنوعی در پیش‌بینی شاخص قیمت سهام و طراحی مدل ترکیبی.
۲. اله یاری، ا. و اکبر. (۱۳۸۷). بررسی شکل ضعیف کارایی بازار سرمایه در بورس اوراق بهادار تهران.
۳. امیری، پ. (۱۳۸۷). مطالعه عوامل حسابداری مؤثر بر شاخص قیمت سهام در بورس اوراق بهادار تهران.
۴. بیگی، س. (۱۳۹۰). مقایسه رویکرد تکنیکی و بنیادی در پیش‌بینی قیمت سهام و ارائه یک مدل ترکیبی با کمک سیستم های هوشمند.
۵. پاکدامن، ع. (۱۳۸۷). بررسی عوامل مؤثر بر تغییر شاخص سهام در بورس اوراق بهادار تهران با رویکرد فازی.
۶. تهرانی، ر. و مرادپور، س. (۱۳۹۱). پیش‌بینی بازده شاخص بورس اوراق بهادار با استفاده از مدل های شبکه عصبی مصنوعی شعاع پایه. *مجله مهندسی مالی و مدیریت اوراق بهادار*.
۷. جوادی، ج. (۱۳۶۹). بررسی تأثیر متغیرهای کلان اقتصادی بر روی شاخص قیمت سهام و تصمیم گیری سرمایه گذاران در بورس اوراق بهادار تهران.
۸. دوانی، غ. (۱۳۸۱). *بورس، سهام و نحوه قیمت گذاری سهام*. تهران: نشر نخستین.
۹. سینایی، ح. و مرتضوی، س. و تیموری اصل، ی. (۱۳۸۱). پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه های عصبی مصنوعی.
۱۰. سینایی، ح. و مرتضوی، س. و تیموری اصل، ی. (۱۳۸۴). پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه های عصبی شعاع پایه. *بررسی های حسابداری و حسابرسی*، ۵۹-۸۳.
۱۱. عباس پور، م. (۱۳۸۱). پیش‌بینی قیمت سهام شرکت ایران خودرو با شبکه عصبی.
۱۲. غفاری، ی. (۱۳۸۳). *راهنمای سرمایه گذاری در بورس*. تبریز: انتشارات شایسه.
۱۳. فدایی نژاد، م. (۱۳۷۴). *بررسی کارایی بورس اوراق بهادار تهران*. تهران.
۱۴. فلاح شمس، ف. و دلنواز، ب. (۱۳۸۸، تابستان). پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه های عصبی. *فراسوی مدیریت*.
۱۵. کیانی، ر. (۱۳۹۰). *نگاهی تحلیلی بر شاخص های بورس*.
۱۶. محمدی تاکامی، س. (۱۳۹۲). *شناسایی و تخمین توابع غیر خطی با شبکه های عصبی و مدل های نروفازی*. تهران: دانشگاه خواجه نصیر الدین طوسی.

۱۷. مشیری، س. و مروت، ح. (۱۳۸۵). پیش‌بینی شاخص کل بازده هی سهام تهران با استفاده از مدل های خطی و غیر خطی.
۱۸. منجمی، س.، ابزری، م. و رعیتی شوارزی، ع. (۱۳۸۸). پیش‌بینی قیمت سهام در بازار بورس اوراق بهادار با استفاده از شبکه ی عصبی فازی والگوریتمهای ژنتیک و مقایسه‌ی آن با شبکه ی عصبی مصنوعی.
۱۹. میرزا خانی، ح. و چوبینه، ب. (۱۳۸۹). جایگاه بورس اوراق بهادار در اقتصاد مقاومتی، سازمان امور اقتصادی و دارایی استان مرکزی.
۲۰. نمازی، م. و محمد مهدی، ک. (۱۳۸۶). پیش‌بینی بازده روزانه سهام شرکت های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از شبکه عصبی مصنوعی. *تحقیقات مالی*، ۱۱۵-۱۳۲.
۲۱. نوروزی موحد، م. (۱۳۹۴). *پیش بینی قیمت سهم در بورس اوراق بهادار به کمک داده کاوی با الگوریتم های ترکیبی تکاملی*. شیراز: دانشگاه شیراز.

22. Harper, R. (2010) "GE, Explosive Grammars and the Lasting Legacy of Bad . (n.d.).
23. O' Neill, M. (2010). Automatic Programming with Grammatical Evolution.
24. O'Neill M., Ryan C. (2001) Grammatical Evolution, IEEE Trans. evolutionary . (n.d.).
25. Ryan C., Collins J.J., O'Neill M. (1998). Grammatical evolution: Evolving Programs for an Arbitrary Language . Lecture Notes in Computer Science 1391, Proceedings of the First European Workshop on Genetic Programming, 83-95, Springer-Verlag. (n.d.).
26. Abderrazak. (2014). day-of the week and returns distribution: evidence from the tunisian stock market.
27. Anderson, M., Banker, R., Huang, R., & Janakirman, S. (2014). cost behavior and fundamental analysis of SG&A cost. *accepting auditing & finance*.
28. Ang, A., & Chen, J. (2014). CAPM over the long run. *Journal of Emprical Finance*.
29. Banwens, L., & Laurent, S. (2015). A new class of multivariate skew densities, with application to generalized outoregressive conditional heteroscedasticity models. *business & economic statistics*.
30. Bernstein, R. (2013). Capital ideas: the improbable origins of modern wall street.
31. Brabazon, A ,O'Neill, M ,Matthews, R & ,Ryan, C .(۲۰۰۶) .Grammatical Evolution and Corporate Failure Prediction .Center for International Business Policy Kingston Business School, London.

32. Chi Lee, M. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction.
33. da silva, p. (2013). modeling and forecasting liquidity using principal component analysis and dynamic factor models.
34. Edwards, R., Bassetti, W., & Magee, J. (2007). Technical analysis of stock trends.
35. Eiben, A., & smith, E. (2013). introduction to Evolutionary computing.
36. G. Jones, "Genetic and Evolutionary Algorithms", University of Sheffield, UK. (n.d.).
37. Goldberg, D.E. 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley . (n.d.).
38. Goldberg, D. (2008). Genetic Algorithms in Search Optimization and Machine Learning.
39. Gount, C. (2004). Size and book to market effect and the Fama French three factor asset pricing model: evidence from the Australian stock market . accounting & finance.
40. Holland, J. (2002). Adaptation in Natural and Artificial Systems. MIT press, Cambridge MA.
41. Kelly, L. (2007). Measuring the Economic Stock of Money.
42. kim, S., & Lee, K. (2011). pricing of liquidity risks by alternative liquidity measure conference.
43. L. J. Fogel, A. J. Owens, and M. J. Walsh, "Artificial Intelligence through Simulated Evolution", Wiley, New York, 1966. (n.d.).
44. Levin, U., & Moody, J. (1993). Fast Pruning Using Principal, Department of Computer Science and Engineering, Oregon Graduate Institute.
45. Lewellen, J., & Negel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. Financial economics, 289-314.
46. Lou, X., & Sho, T. (2014). price impacting or trading volume. Social Science Electronic.
47. mehrazin, & akrami. (2012). the effect of ramadhan month on stock abnormal return of the companies accept in tehran stock exchange. economics and finance review.
48. Mills, T., & Markellos, R. (2010). The economics modelling of the financial time series. Cambridge university press.

49. O'Neill, M. and Ryan, C. (2001) "Grammatical Evolution". IEEE Transactions on Evolutionary Computation 5(4), 349-358. (n.d.).
50. O'Neill, M., & Brabazon, A. (2012). Grammatical Differential Evolution.
51. O'Neill, M., Brabazon, A., & Ryan, C. (2012). Evolving Market Index Trading Rules Using.
52. R, w. (2010). Measuring and predicting Liquidity in the stock market. Gallen.
53. Rechenberg, I. (1973). Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution.
54. Tsang, P., & Kwok, P. (2007). Design and implementation of NN5 for Hong Kong stock price forecasting, Engineering Applications of Artificial Intelligence.
55. W, k., & H, Z. (2013). Measuring liquidity in emerging market. Pacific-Basin Finance.
56. Yakup , K., Melek Acar , B., & mer Kaan, B. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange.
57. Yim, J. (2002). A comparison of neural networks with time series models for forecasting returns on a stock market index. paper. School of Economics and Finance.
58. Zhang, Z. (2005). Stock time series forecasting using support vector machines employing analyst recommendations.
59. Zivot, E., & Wang, J. (2008). Modelling Financial Time series With S-PLUS. Spring Science & Business Media.

پیوست

تابع run 1 برای پیش بینی روز اول بکار گرفته شده است. هر یک از توابعی که توسط run 1 فراخوانی می شود نیز در ادامه آن آورده شده است.

```
clear,clc
tic
rt={'diff point';'iterations';'test
result';'var2';'var10';'var13';'var17';'var18';'fitness'};
%----- Year -1 -----
--
[FRF1]=xlsread('increase indexes, Extra Removed.xlsx','Y -1');%FRF1 is
increase indexes of prior years
[FRS1]=xlsread('decrease indexes, Extra Removed.xlsx','y -1');%FRS1 is
decrease indexes of prior years
dp=100;
[nn,testresult,pp]=priorindex(FRF1,FRS1,dp);
values(:,2)=[dp;nn;testresult;pp'];
xlswrite('Results, Prior index.xlsx',values,'Y -1')
xlswrite('Results, Prior index.xlsx',rt,'Y -1')
toc
```

تابع crossover:

```
function [boy,girl]=crossover(papa,mama)
% we use one points of crossover calling for two points to be selected on
the parent organisms and
% %everything between the one point is swapped between the parents,
rendering two child organisms:
% papa = [2 4 0 7 9];
% mama = [6 2 4 6 0];
% [boy,girl]=crossover(papa,mama)
%boy =[2 4 0 6 0] :D
%girl=[6 2 4 7 9] :P
boy =[papa(1:3),mama(4:5)];
girl=[mama(1:3),papa(4:5)];
```

تابع fitness:

```
function [A]=fitness(gene,FRF,FRS,dp)
Gene(1:10,:)= [gene;gene;gene;gene;gene;gene;gene;gene;gene;gene];
Gene(11:20,:)=Gene;Gene(21:60,:)= [Gene;Gene];
MF=Gene.*FRF(1:60,1:5);%FRF is increase index
MS=Gene.*FRS(1:60,1:5);%FRS is decrease index
Sumf=sum(MF,2);
Sums=sum(MS,2);
A=sum(Sumf<=dp)+sum(Sums>=dp);
```

```

function [nn,testresult,pp]=priorindex(FRF,FRS,dp)
%nn is the number of iterations passed, pp is the best chromosome after
%all, and dp is differentiation point and testresult is the number of
indexes the
%network can predict successfully in test stage
%=====
%=====
%We start with Npop=12 and Nkeep=6 mutation rate of 0.2
%From 80 increase index and decrease index we select 60 from each
%for training.
%===== Initial Population
%=====
P=zeros(12,6);%we preallocate P to stop it from changing size in every
loop
randompool=[-10000*rand,-1000*rand,-100*rand,-
10*rand,0,10*rand,100*rand,1000*rand,10000*rand];
for ii=1:12
    for jj=1:5
        index=round(8*rand)+1;
        randompool(index);
        P(ii,jj)=randompool(index);
    end
end
%===== Assessing initial population and
descending sorting =====
for ii=1:12
    P(ii,6)=fitness(P(ii,1:5),FRF,FRS,dp);
end
P=sortcolumn(P,6,'descend');%sorting P in descending order
%===== Loops
%=====
nn=0;
while P(1,6)<120
    nn=nn+1;
    %----- Selecting competent parents among the first
6 chromosomes -----
    [papa1,mama1]=randparent(P(1:6,:),6,'fitness');
    [papa2,mama2]=randparent(P(1:6,:),6,'fitness');
    [papa3,mama3]=randparent(P(1:6,:),6,'fitness');
    %----- Mating -----
    -----
    [boy1,girl1]=crossover(papa1,mama1);
    [boy2,girl2]=crossover(papa2,mama2);
    [boy3,girl3]=crossover(papa3,mama3);
    %----- Calculating the fitness of
children -----
    boy1(6)=fitness(boy1(1:5),FRF,FRS,dp);girl1(6)=fitness(girl1(1:5),FRF,FRS
,dp);
    boy2(6)=fitness(boy2(1:5),FRF,FRS,dp);girl2(6)=fitness(girl2(1:5),FRF,FRS
,dp);

```



```

boy3(6)=fitness(boy3(1:5),FRF,FRS,dp);girl3(6)=fitness(girl3(1:5),FRF,FRS
,dp);
%----- discarding the unfit and replacing the new
chromosomes -----
ComparingMatrix=[P(7:12,:);boy1;girl1;boy2;girl2;boy3;girl3];
ComparingMatrix=sortcolumn(ComparingMatrix,6,'descend');
P(7:12,:)=ComparingMatrix(1:6,:);
% -----Sorting new entered chromosomes population
-----

P=sortcolumn(P,6,'descend');
%----- Mutation -----
-----

%P is a 12*6 matrice which coefficients are 12*5 i.e. 60
coefficients.
%among these 12 chromosomes, we want three of them who are the top
best
%ones unchanged. we change %10 of the rest coefficients i.e. 9 of
them
%who has the sum of 9*5=45 genes to another random coefficient.
for pp=1:10
    randompool=[-10000*rand,-1000*rand,-100*rand,-
10*rand,0,10*rand,100*rand,1000*rand,10000*rand];
    index=round(8*rand)+1;
    P(3+round(8*rand)+1,round(4*rand)+1)=randompool(index);
end
%----- Calculating the fitness of
mutated children -----
for ii=1:12
    P(ii,6)=fitness(P(ii,1:5),FRF,FRS,dp);
end
%----- Sorting mutated population -
-----

P=sortcolumn(P,6,'descend');
%----- Showing results -----
-----

if mod(nn,10000)==0
    nn
    P(1,6)
end
if nn==100000
    break
end
end
pp=P(1,:);
% ----- Test -----
-----

gene=P(1,1:5);
Gene(1:10,:)=[gene;gene;gene;gene;gene;gene;gene;gene;gene;gene];
Gene(11:20,:)=Gene;
MF=Gene.*FRF(1:20,1:5);%FRF is increase indexes
MS=Gene.*FRS(1:20,1:5);%FRS is decrease indexes
Sumf=sum(MF,2);
Sums=sum(MS,2);
testresult=sum(Sumf<=dp)+sum(Sums>=dp);

```

:randomparent كد

```
function [papa,mama]=randomparent(c,columnno,forc)
%This function is the same roswelt wheel to select the best parent to
%mating. "c" matrix is cromosome matrix and its cromosomes will be
selected
%according to the columnno'th column and they'll be selected whether as a
%fitness(the more the better) or cost(the less the better), which is
%specified by "forc" i.e. fitness or cost.
switch forc
    case {'fitness'}
        rounded=round(c(:,columnno));S=sum(rounded);papano=round((S-
1)*rand)+1;
        for ii=1:size(c,1)
            if (ii==1)&&(papano<=rounded(1))
                papano=1;
                break
            elseif (ii~=1)&&(papano>=sum(rounded([1:ii-
1])))&&(papano<=sum(rounded([1:ii])))
                papano=ii;
                break
            else
                papano=1;
                break
            end
        end
        mamano=round((S-1)*rand)+1;
        for ii=1:size(c,1)
            if (ii==1)&&(mamano<=rounded(1))
                mamano=1;
                break
            elseif (ii~=1)&&(mamano>=sum(rounded([1:ii-
1])))&&(mamano<=sum(rounded([1:ii])))
                mamano=ii;
                break
            else
                mamano=1;
                break
            end
        end
    case {'cost'}

=====
    S=1./c(:,columnno);papano=rand*sum(S);
    for ii=1:size(c,1)
        if (ii==1)&&(papano<=S(1))
            papano=1;
            break
        elseif (ii~=1)&&(papano>=sum(S([1:ii-
1])))&&(papano<=sum(S([1:ii])))
            papano=ii;
            break
        else
```

```

        papano=1;
        break
    end
end
mamano=rand*sum(S);
for ii=1:size(c,1)
    if (ii==1) && (mamano<=S(1))
        mamano=1;
        break
    elseif (ii~=1) && (mamano>=sum(S([1:ii-1]))) && (mamano<=sum(S([1:ii])))
        mamano=ii;
        break
    else
        mamano=1;
        break
    end
end
end
papa=c(papano,:);mama=c(mamano,:);
%c=[7 4 3 6 8;-6 5 2 3 7;8 4 0 20 4]
%[papa,mama]=randparent(c,5,'cost')
%[papa,mama]=randparent(c,5,'fitness')

%c=[7 4 3 6 8;-6 5 2 3 7;8 4 0 20 4;1000,1,1,3,4]
%[papa,mama]=randparent(c,6,'cost')
%[papa,mama]=randparent(c,6,'fitness')

```

تابع sortcolumn:

```

function c=sortcolumn(a,columnno,b)
narginchk(2,3)
if nargin==2
    b='ascend';
end
if columnno>size(a,2) || columnno<1
    error('Your no of column is invalid')
end
switch b
    case{'ascend'}
        n=1;
        while n~=size(a,1);
            for ii=1:size(a,1)-1;
                if a(ii,columnno)>a(ii+1,columnno);
                    r=a(ii,:);
                    a(ii,:)=a(ii+1,:);
                    a(ii+1,:)=r;
                end
            end
            n=1;
            for ii=1:size(a,1)-1;
                if a(ii,columnno)<=a(ii+1,columnno);
                    n=n+1;
                end
            end
        end
    end
end

```

```

        end
    end
    case{ 'descend' }
    n=1;
    while n~=size(a,1);
        for ii=1:size(a,1)-1;
            if a(ii,columnno)<a(ii+1,columnno);
                r=a(ii,:);
                a(ii,:)=a(ii+1,:);
                a(ii+1,:)=r;
            end
        end
        n=1;
        for ii=1:size(a,1)-1;
            if a(ii,columnno)>=a(ii+1,columnno);
                n=n+1;
            end
        end
    end
end
c=a;
%a=[2 8 -2; 6 0 1; -2 5 9; 2 4 8];sortcolumn(a,2,'descend')
%sorts matrix a vs it's 2nd column in a descending order

```

تابع run 2 برای پیش بینی روز دوم بکار گرفته شده است. هر یک از توابعی که توسط run 2 فراخوانی می شود نیز در ادامه آن آورده شده است.

```

clear,clc
tic
rt={'diff point';'iterations';'test
result';'var3';'var10';'var16';'var17';'var18';'fitness'};
%----- Year -2 -----
--
[FRF2]=xlsread('increase indexes, Extra Removed.xlsx','Y -2');%FRF2 is
increase indexes of prior years
[FRS2]=xlsread('decrease indexes, Extra Removed.xlsx','y -2');%FRS2 is
decrease indexes of prior years
dp=300;
[nn,testresult,pp]=prioryears(FRF2,FRS2,dp);-
values(:,2)=[dp;nn;testresult;pp'];
xlswrite('Results, Prior index.xlsx',values,'Y -2')
xlswrite('Results, Prior index.xlsx',rt,'Y -2')
toc

```

تابع crossover:

```

function [boy,girl]=crossover(papa,mama)
% we use one points of crossover calling for two points to be selected on
the parent organisms and
% everything between the one point is swapped between the parents,
rendering two child organisms:

```

```

% papa = [2 4 0 7 9];
% mama = [6 2 4 6 0];
% [boy,girl]=crossover(papa,mama)
%boy =[2 4 0 6 0] :D
%girl=[6 2 4 7 9] :P
boy =[papa(1:3),mama(4:5)];
girl=[mama(1:3),papa(4:5)];

```

تابع fitness:

```

function [A]=fitness (gene,FRF,FRS,dp)
Gene(1:10,:)= [gene;gene;gene;gene;gene;gene;gene;gene;gene;gene];
Gene(11:20,:)=Gene;Gene(21:60,:)= [Gene;Gene];
MF=Gene.*FRF(1:60,1:5);%FRF is increase index
MS=Gene.*FRS(1:60,1:5);%FRS is decrease index
Sumf=sum(MF,2);
Sums=sum(MS,2);
A=sum(Sumf<=dp)+sum(Sums>=dp);

```

تابع priorindex:

```

function [nn,testresult,pp]=priorindex(FRF,FRS,dp)
%nn is the number of iterations passed, pp is the best chromosome after
%all, and dp is differentiation point and testresult is the number of
%firms the
%network can predict successfully in test stage
%=====
%=====
%We start with Npop=12 and Nkeep=6 mutation rate of 0.2
%From 80 increase index and 60 decrease index we select 60 from each
%for training.
%===== Initial Population
%=====
P=zeros(12,6);%we preallocate P to stop it from changing size in every
loop
randpool=[-10000*rand,-1000*rand,-100*rand,-
10*rand,0,10*rand,100*rand,1000*rand,10000*rand];
for ii=1:12
    for jj=1:5
        index=round(8*rand)+1;
        randpool(index);
        P(ii,jj)=randpool(index);
    end
end
%===== Assessing initial population and
descending sorting =====
for ii=1:12
    P(ii,6)=fitness(P(ii,1:5),FRF,FRS,dp);
end
P=sortcolumn(P,6,'descend');%sorting P in descending order

```

```

%===== Loops
=====
nn=0;
while P(1,6)<120
    nn=nn+1;
    %----- Selecting competent parents among the first
6 chromosomes -----
    [papa1,mama1]=randomparent(P(1:6,:),6,'fitness');
    [papa2,mama2]=randomparent(P(1:6,:),6,'fitness');
    [papa3,mama3]=randomparent(P(1:6,:),6,'fitness');
    %----- Mating -----
    [boy1,girl1]=crossover(papa1,mama1);
    [boy2,girl2]=crossover(papa2,mama2);
    [boy3,girl3]=crossover(papa3,mama3);
    %----- Calculating the fitness of
children -----
    boy1(6)=fitness(boy1(1:5),FRF,FRS,dp);girl1(6)=fitness(girl1(1:5),FRF,FRS
,dp);
    boy2(6)=fitness(boy2(1:5),FRF,FRS,dp);girl2(6)=fitness(girl2(1:5),FRF,FRS
,dp);
    boy3(6)=fitness(boy3(1:5),FRF,FRS,dp);girl3(6)=fitness(girl3(1:5),FRF,FRS
,dp);
    %----- discarding the unfit and replacing the new
chromosomes -----
    ComparingMatrix=[P(7:12,:);boy1;girl1;boy2;girl2;boy3;girl3];
    ComparingMatrix=sortcolumn(ComparingMatrix,6,'descend');
    P(7:12,:)=ComparingMatrix(1:6,:);
    % -----Sorting new entered chromosomes population
-----
    P=sortcolumn(P,6,'descend');
    %----- Mutation -----
    %P is a 12*6 matrice which coefficients are 12*5 i.e. 60
coefficients.
    %among these 12 chromosomes, we want three of them who are the top
best
    %ones unchanged. we change %10 of the rest coefficients i.e. 9 of
them
    %who has the sum of 9*5=45 genes to another random coefficient.
    for pp=1:10
        randompool=[-10000*rand,-1000*rand,-100*rand,-
10*rand,0,10*rand,100*rand,1000*rand,10000*rand];
        index=round(8*rand)+1;
        P(3+round(8*rand)+1,round(4*rand)+1)=randompool(index);
    end
    %----- Calculating the fitness of
mutated children -----
    for ii=1:12
        P(ii,6)=fitness(P(ii,1:5),FRF,FRS,dp);
    end
    %----- Sorting mutated population -
-----

```

```

P=sortcolumn(P,6,'descend');
%----- Showing results -----
-----
if mod(nn,10000)==0
    nn
    P(1,6)
end
if nn==100000
    break
end
end
pp=P(1,:);
% ----- Test -----
-----
gene=P(1,1:5);
Gene(1:10,:)=[gene;gene;gene;gene;gene;gene;gene;gene;gene;gene];
Gene(11:20,:)=Gene;
MF=Gene.*FRF(1:20,1:5);%FRF is increase index
MS=Gene.*FRS(1:20,1:5);%FRS is decrease index
Sumf=sum(MF,2);
Sums=sum(MS,2);
testresult=sum(Sumf<=dp)+sum(Sums>=dp);

```

فایل :rundomparent

```

function [papa,mama]=rundomparent(c,columnno,forc)
%This function is the same roswelt wheel to select the best parent to
% mating. "c" matrix is cromosome matrix and its cromosomes will be
% selected
% according to the columnno'th column and they'll be selected whether as a
% fitness(the more the better) or cost(the less the better), which is
% specified by "forc" i.e. fitness or cost.
switch forc
    case {'fitness'}
        rounded=round(c(:,columnno));S=sum(rounded);papano=round((S-
1)*rand)+1;
        for ii=1:size(c,1)
            if (ii==1)&&(papano<=rounded(1))
                papano=1;
                break
            elseif (ii~=1)&&(papano>=sum(rounded([1:ii-
1])))&&(papano<=sum(rounded([1:ii])))
                papano=ii;
                break
            else
                papano=1;
                break
            end
        end
        mamano=round((S-1)*rand)+1;
        for ii=1:size(c,1)

```

```

        if (ii==1) && (mamano<=rounded(1))
            mamano=1;
            break
        elseif (ii~=1) && (mamano>=sum(rounded([1:ii-
1]))) && (mamano<=sum(rounded([1:ii])))
            mamano=ii;
            break
        else
            mamano=1;
            break
        end
    end
end
case {'cost'}

%=====
S=1./c(:,columnno);papano=rand*sum(S);
for ii=1:size(c,1)
    if (ii==1) && (papano<=S(1))
        papano=1;
        break
    elseif (ii~=1) && (papano>=sum(S([1:ii-
1]))) && (papano<=sum(S([1:ii])))
        papano=ii;
        break
    else
        papano=1;
        break
    end
end
mamano=rand*sum(S);
for ii=1:size(c,1)
    if (ii==1) && (mamano<=S(1))
        mamano=1;
        break
    elseif (ii~=1) && (mamano>=sum(S([1:ii-
1]))) && (mamano<=sum(S([1:ii])))
        mamano=ii;
        break
    else
        mamano=1;
        break
    end
end
end
papa=c(papano,:);mama=c(mamano,:);
%c=[7 4 3 6 8;-6 5 2 3 7;8 4 0 20 4]
%[papa,mama]=randparent(c,5,'cost')
%[papa,mama]=randparent(c,5,'fitness')

%c=[7 4 3 6 8;-6 5 2 3 7;8 4 0 20 4;1000,1,1,3,4]
%[papa,mama]=randparent(c,6,'cost')
%[papa,mama]=randparent(c,6,'fitness')

```



```

function c=sortcolumn(a,columnno,b)
narginchk(2,3)
if nargin==2
    b='ascend';
end
if columnno>size(a,2)||columnno<1
    error('Your no of column is invalid')
end
switch b
    case{'ascend'}
        n=1;
        while n~=size(a,1);
            for ii=1:size(a,1)-1;
                if a(ii,columnno)>a(ii+1,columnno);
                    r=a(ii,:);
                    a(ii,:)=a(ii+1,:);
                    a(ii+1,:)=r;
                end
            end
            n=1;
            for ii=1:size(a,1)-1;
                if a(ii,columnno)<=a(ii+1,columnno);
                    n=n+1;
                end
            end
        end
    case{'descend'}
        n=1;
        while n~=size(a,1);
            for ii=1:size(a,1)-1;
                if a(ii,columnno)<a(ii+1,columnno);
                    r=a(ii,:);
                    a(ii,:)=a(ii+1,:);
                    a(ii+1,:)=r;
                end
            end
            n=1;
            for ii=1:size(a,1)-1;
                if a(ii,columnno)>=a(ii+1,columnno);
                    n=n+1;
                end
            end
        end
    end
end
c=a;
%a=[2 8 -2; 6 0 1; -2 5 9; 2 4 8];sortcolumn(a,2,'descend')
%sorts matrix a vs it's 2nd column in a descending order

```

تابع run 3 برای پیش بینی روز سوم بکار گرفته شده است. هر یک از توابعی که توسط run 3 فراخوانی می شود نیز در ادامه آن آورده شده است.

```

Clear,clc
tic
rt={'diff point';'iterations';'test
result';'var1';'var4';'var14';'var15';'var18';'fitness'};
%----- Year -3 -----
--
[FRF3]=xlsread('increase index, Extra Removed.xlsx','Y -3');%FRF3 is
increase index
[FRS3]=xlsread('decrease index, Extra Removed.xlsx','y -3');%FRS3 is
decrease index
dp=300;
[nn,testresult,pp]=prioryears(FRF3,FRS3,dp);
values(:,2)=[dp;nn;testresult;pp'];
xlswrite('Results, Prior index.xlsx',values,'Y -3')
xlswrite('Results, Prior index.xlsx',rt,'Y -3')
testresult
toc

```

تابع crossover:

```

function [boy,girl]=crossover(papa,mama)
% we use one points of crossover calling for two points to be selected on
the parent organisms and
% %everything between the one point is swapped between the parents,
rendering two child organisms:
% papa = [2 4 0 7 9];
% mama = [6 2 4 6 0];
% [boy,girl]=crossover(papa,mama)
%boy =[2 4 0 6 0] :D
%girl=[6 2 4 7 9] :P
boy =[papa(1:3),mama(4:5)];
girl=[mama(1:3),papa(4:5)];

```

تابع fitness:

```

function [A]=fitness(gene,FRF,FRS,dp)
Gene(1:10,⊕)=[gene;gene;gene;gene;gene;gene;gene;gene;gene;gene];
Gene(11:20,⊕)=Gene;Gene(21:60,⊕)=[Gene;Gene];
MF=Gene.*FRF(1:60,1:5);%FRF is increase index
MS=Gene.*FRS(1:60,1:5);%FRS is decrease index
Sumf=sum(MF,2);
Sums=sum(MS,2);
A=sum(Sumf<=dp)+sum(Sums>=dp);

```

priorindex ك

```
function [nn,testresult,pp]=priorindex(FRF,FRS,dp)
%nn is the number of iterations passed, pp is the best chromosome after
%all, and dp is differentiation point and testresult is the number of
%network can predict successfully in test stage
%=====
%We start with Npop=12 and Nkeep=6 mutation rate of 0.2
%From 80 increase index and 60 decrease index we select 60 from each
%for training.
%===== Initial Population
P=zeros(12,6);%we preallocate P to stop it from changing size in every
loop
randompool=[-10000*rand,-1000*rand,-100*rand,-
10*rand,0,10*rand,100*rand,1000*rand,10000*rand];
for ii=1:12
    for jj=1:5
        index=round(8*rand)+1;
        randompool(index);
        P(ii,jj)=randompool(index);
    end
end
%===== Assessing initial population and
descending sorting =====
for ii=1:12
    P(ii,6)=fitness(P(ii,1:5),FRF,FRS,dp);
end
P=sortcolumn(P,6,'descend');%sorting P in descending order
%===== Loops
nn=0;
while P(1,6)<120
    nn=nn+1;
    %----- Selecting competent parents among the first
6 chromosomes -----
    [papa1,mama1]=randomparent(P(1:6,:),6,'fitness');
    [papa2,mama2]=randomparent(P(1:6,:),6,'fitness');
    [papa3,mama3]=randomparent(P(1:6,:),6,'fitness');
    %----- Mating -----
    [boy1,girl1]=crossover(papa1,mama1);
    [boy2,girl2]=crossover(papa2,mama2);
    [boy3,girl3]=crossover(papa3,mama3);
    %----- Calculating the fitness of
children -----
    boy1(6)=fitness(boy1(1:5),FRF,FRS,dp);girl1(6)=fitness(girl1(1:5),FRF,FRS
,dp);
    boy2(6)=fitness(boy2(1:5),FRF,FRS,dp);girl2(6)=fitness(girl2(1:5),FRF,FRS
,dp);
```

```

boy3(6)=fitness(boy3(1:5),FRF,FRS,dp);girl3(6)=fitness(girl3(1:5),FRF,FRS
,dp);
%----- discarding the unfit and replacing the new
chromosomes -----
ComparingMatrix=[P(7:12,:);boy1;girl1;boy2;girl2;boy3;girl3];
ComparingMatrix=sortcolumn(ComparingMatrix,6,'descend');
P(7:12,:)=ComparingMatrix(1:6,:);
% -----Sorting new entered chromosomes population
-----
P=sortcolumn(P,6,'descend');
%----- Mutation -----
-----
%P is a 12*6 matrice which coefficients are 12*5 i.e. 60
coefficients.
%among these 12 chromosomes, we want three of them who are the top
best
%ones unchanged. we change %10 of the rest coefficients i.e. 9 of
them
%who has the sum of 9*5=45 genes to another random coefficient.
for pp=1:10
    randompool=[-10000*rand,-1000*rand,-100*rand,-
10*rand,0,10*rand,100*rand,1000*rand,10000*rand];
    index=round(8*rand)+1;
    P(3+round(8*rand)+1,round(4*rand)+1)=randompool(index);
end
%----- Calculating the fitness of
mutated children -----
for ii=1:12
    P(ii,6)=fitness(P(ii,1:5),FRF,FRS,dp);
end
%----- Sorting mutated population -
-----
P=sortcolumn(P,6,'descend');
%----- Showing results -----
-----
if mod(nn,10000)==0
    nn
    P(1,6)
end
if nn==100000
    break
end
end
pp=P(1,:)
% ----- Test -----
-----
gene=P(1,1:5);
Gene(1:10,:)=[gene;gene;gene;gene;gene;gene;gene;gene;gene;gene];
Gene(11:20,:)=Gene;
MF=Gene.*FRF(1:20,1:5);%FRF is increase index
MS=Gene.*FRS(1:20,1:5);%FRS is decrease index
Sumf=sum(MF,2);
Sums=sum(MS,2);
testresult=sum(Sumf<=dp)+sum(Sums>=dp);

```

```

function [papa,mama]=randomparent(c,columnno,forc)
%This function is the same roswelt wheel to select the best parent to
% mating. "c" matrix is cromosome matrix and its cromosomes will be
% selected
% according to the columnno'th column and they'll be selected whether as a
% fitness(the more the better) or cost(the less the better), which is
% specified by "forc" i.e. fitness or cost.
switch forc
    case {'fitness'}
        rounded=round(c(:,columnno));S=sum(rounded);papano=round((S-
1)*rand)+1;
        for ii=1:size(c,1)
            if (ii==1)&&(papano<=rounded(1))
                papano=1;
                break
            elseif (ii~=1)&&(papano>=sum(rounded([1:ii-
1])))&&(papano<=sum(rounded([1:ii])))
                papano=ii;
                break
            else
                papano=1;
                break
            end
        end
        mamano=round((S-1)*rand)+1;
        for ii=1:size(c,1)
            if (ii==1)&&(mamano<=rounded(1))
                mamano=1;
                break
            elseif (ii~=1)&&(mamano>=sum(rounded([1:ii-
1])))&&(mamano<=sum(rounded([1:ii])))
                mamano=ii;
                break
            else
                mamano=1;
                break
            end
        end
    case {'cost'}

%=====
        S=1./c(:,columnno);papano=rand*sum(S);
        for ii=1:size(c,1)
            if (ii==1)&&(papano<=S(1))
                papano=1;
                break
            elseif (ii~=1)&&(papano>=sum(S([1:ii-
1])))&&(papano<=sum(S([1:ii])))
                papano=ii;
                break
            else
                papano=1;
                break
            end
        end
    end
end

```

```

        end
    end
    mamano=rand*sum(S);
    for ii=1:size(c,1)
        if (ii==1) && (mamano<=S(1))
            mamano=1;
            break
        elseif (ii~=1) && (mamano>=sum(S([1:ii-1]))) && (mamano<=sum(S([1:ii])))
            mamano=ii;
            break
        else
            mamano=1;
            break
        end
    end
end
end
papa=c(papano,:);mama=c(mamano,:);
%c=[7 4 3 6 8;-6 5 2 3 7;8 4 0 20 4]
%[papa,mama]=randparent(c,5,'cost')
%[papa,mama]=randparent(c,5,'fitness')

%c=[7 4 3 6 8;-6 5 2 3 7;8 4 0 20 4;1000,1,1,3,4]
%[papa,mama]=randparent(c,6,'cost')
%[papa,mama]=randparent(c,6,'fitness')

```

تابع sortcolumn :

```

function c=sortcolumn(a,columnno,b)
narginchk(2,3)
if nargin==2
    b='ascend';
end
if columnno>size(a,2)||columnno<1
    error('Your no of column is invalid')
end
switch b
    case{'ascend'}
        n=1;
        while n~=size(a,1);
            for ii=1:size(a,1)-1;
                if a(ii,columnno)>a(ii+1,columnno);
                    r=a(ii,:);
                    a(ii,:)=a(ii+1,:);
                    a(ii+1,:)=r;
                end
            end
            n=n+1;
        end
    case{'descend'}
        n=1;
        for ii=1:size(a,1)-1;
            if a(ii,columnno)<=a(ii+1,columnno);
                n=n+1;
            end
        end
end
end

```

```

case{'descend'}
n=1;
while n~=size(a,1);
    for ii=1:size(a,1)-1;
        if a(ii,columnno)<a(ii+1,columnno);
            r=a(ii,:);
            a(ii,:)=a(ii+1,:);
            a(ii+1,:)=r;
        end
    end
    n=1;
    for ii=1:size(a,1)-1;
        if a(ii,columnno)>=a(ii+1,columnno);
            n=n+1;
        end
    end
end
end
end
c=a;
%a=[2 8 -2; 6 0 1; -2 5 9; 2 4 8];sortcolumn(a,2,'descend')
%sorts matrix a vs it's 2nd column in a descending order

```


Abstract

A lot of research to predict and reliable in the field Stock Exchange index has been forecast. Research from 2008 to November 2016 the time domain and the spatial domain Tehran Stock Exchange.

This research tries to use the grammatical evolution algorithm to predict stock index securities. Former grammatical evolution algorithm to explore the potential for useful rules for predicting bankruptcy of the Company solely in the field of applied and been fruitful. Here are the total index and other related indices, the daily and during the last 8 years that records have been used. Results show that the model used to predict the total index is acceptable capability.

Keywords: Index prediction models, Grammatical Evolution algorithm, Forecasting Index



Shahrood University of Technology
Faculty of Industrial Engineering and Management
M.Sc Thesis in Master of Business Administration

**Forecasting Stock Exchange index using grammatical evolution
algorithm**

By: Mohammadali Yarahmadi

Supervisor:

Dr. Seyyed Mojtba Mirlohi

Advisor:

Dr. Aliakbar Hasani

February 2017