دوفصلنامه اقتصاد پولی، مالی(دانش وتوسعه سابق) دوره جدید، سال بیستم، شمارهٔ۶، پاییز و زمستان ۱۳۹۲

کاربرد روش k- نزدیک‌ترین همسایه در پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران

محمود موسوی شیری[۱]
استادیار گروه حسابداری، دانشگاه پیام نور
صادق بافنده ایماندوست[*]
استادیار گروه اقتصاد، دانشگاه پیام نور،
محمد بلندرفتار پسیخانی [۲]
کارشناس ارشد مدیریت، دانشگاه پیام نور

تاریخ دریافت : ۱۳۹۱/۱۰/۱۸ **تاریخ پذیرش:** ۱۳۹۲/۳/۲۷

چکیده

با توجه به تاثیرات درماندگی مالی شرکت‌ها بر گروه‌های ذینفع، همواره ارائه الگوهای پیش‌بینی درماندگی مالی یکی از جذاب‌ترین حوزه‌ها در تحقیقات مالی بوده است. در سال‌های اخیر و پس از رخ دادن بحران مالی جهانی، تعداد شرکت‌های ورشکسته افزایش یافته است. از آن‌جا که درماندگی مالی شرکت‌ها مقدمه ورشکستگی آنهاست، لذا استفاده از نسبت‌های مالی برای پیش‌بینی درماندگی مالی، مورد توجه هر چه بیشتر دانشگاهیان و هم‌چنین بنگاه‌های اقتصادی و نهادهای مالی قرار گرفته است.

گر چه، در سال‌های اخیر در کشور ما نیز تحقیقات انجام شده بر روی پیش‌بینی درماندگی مالی شرکت‌ها رو به افزایش است، با این حال بیشتر تلاش‌های صورت گرفته به استفاده از روش‌های سنتی آماری معطوف شده و تنها در تعداد کمی از پژوهش‌ها از روش‌های ناپارامتری استفاده شده است. نتایج تحقیقات انجام شده در سال‌های اخیر حاکی از آن است که روش‌های یادگیری ماشینی[۱] عملکرد بهتری را نسبت به

---

* - نویسنده مسئول: imandoust@pnu.ac.ir

1- mousavi1973@yahoo.com
2- m.bolandraftar@gmail.com

۱- Machine Learning

روش‌های سنتی آماری دارا می‌باشند.
در این پژوهش از روش k نزدیک‌ترین همسایه برای پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۸۶-۱۳۸۴ استفاده شده است. نتایج حاکی از آن است که روش مذکور با دقت خوبی قادر به پیش‌بینی درماندگی مالی شرکت‌ها می‌باشد.

واژه‌های کلیدی: پیش‌بینی درماندگی مالی شرکت‌ها، یادگیری ماشینی، k-نزدیک‌ترین همسایه، بورس اوراق بهادار تهران
طبقه بندی JEL: M4, G17, G33

Application of K-Nearest Neighbor (KNN) for Predicting Corporate Financial Distress in Tehran Stock Exchange

Mahmoud Mousavi Shiri
Assistant professor of Accounting, Payame Noor University,
Sadegh Bafandeh Imandoust
Assistant professor of Economics, Payame Noor University,
Mohammad Bolandraftar Pasikhani
M.A in Management, Payame Noor University

Received: 7 Jan 2013 **Accept:** 17 Jun 2013

Abstract

Due to the effects of companies' financial distress on stakeholders, financial distress prediction models have been one of the most attractive scopes in financial research. In recent years, after the global financial crisis, the number of bankrupt companies has risen. Since companies' financial distress is the first stage of bankruptcy, using financial ratios for predicting financial distress have attracted too much attention of the academics as well as economic and financial institutions.
Although in recent years studies on predicting companies' financial distress in Iran have been increased, but most efforts have exploited traditional statistical methods; and just a few studies have used nonparametric methods. Recent studies demonstrate machine learning techniques outperform traditional statistical methods.
In the present study *k*-Nearest Neighbor classification method, derived from the field of data mining, is employed to predict financial distress of Tehran Stock Exchange listed companies during 2005-2008. Experimental results show that *k*-Nearest Neighbor is able to predict corporate financial distress with high accuracy.

Keywords: Predicting financial distress, Machine learning, *k*-Nearest Neighbor, Tehran Stock Exchange (TSE)

JEL classification: M4, G17, G33

۱. مقدمه

دانشمندان علوم مالی و اقتصادی برای مطالعه وپیش‌بینی وضعیت مالی بنگاه‌های اقتصادی از روش‌های مختلفی استفاده می‌کنند. امروزه با جهانی‌سازی اقتصاد، رقابت به یکی از مکانیزم‌های اصلی بازار تبدیل شده است. کسب و کارهایی که مدیریت ناکارا و عدم نوآوری داشته باشند، در چنین محیط رقابتی که سرشار از عدم اطمینان است، از کسب پیروزی باز خواهند ماند. در نتیجه، هر بنگاهی در معرض بحران‌های مختلفی مانند بحران بازاریابی، بحران منابع انسانی، بحران کسب اعتبار[1] و بحران نوآوری قرار گیرد، می‌تواند به درماندگی مالی و حتی ورشکستگی مبتلا شود (Sin & Li, 2009).

در چنین محیطی، سرمایه‌گذاران همواره می‌خواهند با پیش‌بینی امکان ورشکستگی یک شرکت، از ریسک سوخت شدن اصل و فرع سرمایه خود جلوگیری کنند. از این‌رو، در پی روش‌هایی هستند که بتوانند به وسیله آن ورشکستگی مالی شرکت‌ها را تخمین بزنند، زیرا در صورت ورشکستگی، قیمت سهام شرکت‌ها به شدت کاهش می‌یابد (Rasoolzadeh, 2002).

با توجه به رابطه میان درماندگی مالی و ورشکستگی شرکت، این نوع درماندگی قابل توجه‌ترین نوع درماندگی برای شرکت‌ها محسوب می‌شود و نه تنها صدمات مالی زیادی بر شرکت‌ها وارد می‌کند، بلکه بر بقا و توسعه آنها نیز تاثیر مستقیم دارد. وانگهی، هنگامی‌که در یک دوره زمانی یکسان شرکت‌های زیادی دچار درماندگی مالی شوند، ممکن است این امر سبب ایجاد بحران مالی در بازار سرمایه گردد. از این‌رو، یافتن راهکارهای اساسی برای کمک به شرکت‌ها، به موضوع داغی برای تحقیق، هم در محیط دانشگاهی و هم در دنیای واقعی، تبدیل شده است (Sin & Li, 2009). با توجه به افزایش روز افزون تعداد شرکت‌های درمانده مالی که افزایش شرکت‌های ورشکسته را به دنبال خواهد داشت، نیاز به پیش‌بینی آینده مالی شرکت‌ها، هر روز بیشتر احساس می‌شود. در نتیجه، ارائه مدل‌های پیش‌بینی درماندگی مالی برای ذینفعان سازمان‌ها شامل میزان اعتباردهندگان و سهامداران، در تصمیم‌گیری‌ها بسیار سودمند خواهد بود.

1- Crediting Crisis

پیش‌بینی درماندگی مالی با استفاده از روش‌های مختلفی صورت می‌پذیرد که از میان آنها، روش تجزیه و تحلیل نسبت‌ها و روش تجزیه و تحلیل ریسک بازار از اعتبار بیشتری برخوردار است. در روش تجزیه و تحلیل ریسک بازار، احتمال وقوع ورشکستگی شرکت از طریق تغییراتی که در ریسک بازار (مثل واریانس نرخ بازده یک سهم و ریسک سیستماتیک) رخ می‌دهد، تخمین زده می‌شود و در روش تجزیه و تحلیل نسبت‌ها، احتمال وقوع ورشکستگی به وسیله یک گروه از نسبت‌های مالی تخمین زده می‌شود (Rasoolzadeh, 2002).

از آنجا که در مدل‌های پیش‌بینی درماندگی مالی شرکت‌ها، متغیر وابسته از نوع قطعی (درمانده مالی یا سالم) می‌باشد، لذا با مساله‌ای از نوع دسته‌بندی روبرو هستیم. بنابراین، مشخص است که در چنین مطالعاتی می بایست از مدل‌های آماری سنتی مانند تحلیل ممیزی چندگانه[1]، تحلیل لوجیت[2] و تحلیل پروبیت[3] استفاده شود. البته قدرت این روش‌های آماری به برخی فرضیات محدودکننده مانند خطی بودن، نرمال بودن، مستقل‌بودن متغیرهای پیش‌بین از یکدیگر و ... بستگی زیادی دارد (Zhang, 1999). بر خلاف روش‌های سنتی آماری، روش‌های داده‌کاوی مانند ماشین‌های بردار پشتیبان، شبکه‌های عصبی مصنوعی، k-نزدیک‌ترین همسایه و روش بیزی ساده، با توجه به ویژگی‌های غیر خطی و ناپارامتری که دارند، ابزار قدرتمندی در زمینه شناسایی و دسته-بندی الگوها هستند.

پژوهش حاضر که از روش داده‌کاوی k-نزدیک‌ترین همسایه[4] و با روش تجزیه و تحلیل نسبت‌ها برای پیش‌بینی درماندگی مالی شرکت‌ها استفاده کرده است، می‌تواند برای طراحی سیستم امتیازدهی اعتباری در موسسات مالی مورد استفاده قرار گیرد.

۲. پیشینه پژوهش‌های انجام شده

با توجه به ادبیات موضوعی و پیشینه پژوهش‌های گذشته در ارتباط با پیش‌بینی ورشکستگی

1- Multiple Discriminant Analysis (MDA)
2- Logit
3- Probit
4- K-Nearest Neighbor (KNN)

شرکت‌ها، دو طبقه کلی از مدل‌های پیش‌بینی مبتنی بر نسبت‌های مالی مشهود است:

مطالعاتی که با بررسی مجزای تک تک نسبت‌ها اقدام به پیش‌بینی می‌کنند (تحلیل تک متغیره).

مطالعاتی که با استفاده از چندین نسبت و وزن دادن به آنها برای رسیدن به تابع پیش‌بینی ورشکستگی تلاش می‌کنند (تحلیل‌های چند متغیره). این مطالعات خود به دو دسته که هر یک چندین زیرشاخه دارد، قابل تقسیم است، و شامل مدل‌های پارامتریک و مدل‌های ناپارامتری است. مدل‌های تحلیل تمایزی و احتمالات شرطی (پروبیت و لوجیت) زیر شاخه‌های مدل پارامتریک و مدل‌های داده‌کاوی زیر شاخه‌های مدل ناپارامتری محسوب می‌شوند.

بیور[1] (۱۹۶۶) آزمون طبقه‌بندی دوگانه[2] را برای ایجاد مدل پیش‌بینی تک متغیره ورشکستگی انتخاب کرد. او در این پژوهش، عدم‌توانایی شرکت به انجام تعهدات مالی‌اش را به عنوان درماندگی مالی تعریف کرد. در تحقیق مزبور، بیور ۳۰ نسبت مالی را که تصور می‌کرد بهترین شاخص برای سلامت مالی یک شرکت هستند، انتخاب کرد. سپس این نسبت‌ها را به شش گروه تقسیم کرد. این شش گروه عبارت بودند از: نسبت‌های مربوط به جریان نقدی، نسبت‌های بدهی به کل دارایی‌ها، نسبت‌های دارایی‌های نقد شونده به کل دارایی‌ها، نسبت‌های دارایی‌های نقد شونده به بدهی‌های جاری، نسبت‌های فعالیت (گردش) و نسبت‌های سود خالص.

در تحقیقات بعدی مدل‌های چند متغیره مورد توجه قرار گرفتند. در واقع، محققان دریافته بودند که عوامل گوناگونی وجود دارند که به طور همزمان بر درماندگی مالی شرکت‌ها اثر دارند.

ادوارد آلتمن[3] (۱۹۶۸) برای اولین بار اثر ترکیبات مختلف نسبت‌های مالی را برای پیش‌بینی درماندگی مالی شرکت‌ها بررسی کرد. آلتمن تحلیل تک متغیره را به سبب ارایه نتایج گیج‌کننده و متضاد با معیارهای تفکیک‌کننده مورد انتقاد قرار داده و تحلیل تمایزی چندگانه[4] را که در آن چندین نسبت مالی به طور همزمان در پیش‌بینی ورشکستگی بررسی می‌شدند برگزید. مدلی که او

1- Beaver
2- Dichotomous classification test
3- Edward Altman
4- Multiple Discriminant Analysis

به‌دست آورد و به «Z-Score» معروف است، هنوز به عنوان شاخصی برای سلامت مالی شرکت‌ها مورد استفاده قرار می‌گیرد. تئوری اصلی آلتمن این بود که مدل پیش‌بینی ورشکستگی او که از ۵ نسبت مالی تشکیل می شود، می‌تواند برای تشخیص شرکت‌های ورشکسته و سالم مورد استفاده قرار گیرد.

اوهلسون[1] (۱۹۸۰) مدلی را با استفاده از تکنیک لوجیت[2] توسعه داد. وی در پژوهش مذکور از ۱۰۵ شرکت ورشکسته و ۲۰۵۸ شرکت غیر ورشکسته بین سال‌های ۱۹۷۰ تا ۱۹۷۶ استفاده کرد. او نه نسبت مالی را به عنوان متغیر مستقل به کار برد. از این نه نسبت مالی، پنج‌تای آنها در مطالعات قبلی نیز استفاده شده بودند.

در سال‌های اخیر و با پیشرفت‌های صورت گرفته در علوم کامپیوتر، استفاده از مدل‌های داده-کاوی رو به افزایش است.

از اواخر دهه ۱۹۸۰ میلادی تکنیک‌های هوش مصنوعی و یادگیری ماشینی مانند شبکه‌های عصبی مصنوعی به‌طور موفقی برای پیش‌بینی ورشکستگی و درماندگی مالی مورد استفاده قرار گرفته‌اند (Markham,1995 ; Jensen,1992; Malhotra,2002; Elmer,1988; Ragsdale,1995).

در اواخر دهه ۱۹۹۰ میلادی روش ماشین‌های بردار پشتیبان برای حل مسائل دسته‌بندی عرضه گردید.

فن[3] و پالانیسوامی[4] (۲۰۰۰) از این روش برای انتخاب متغیرهای پیش‌بین درماندگی مالی استفاده کردند.

مین[5] و لی[6] (۲۰۰۵) نیز از ماشین‌های بردار پشتیبان برای پیش‌بینی ورشکستگی شرکت‌ها استفاده کردند.

1- Ohlson
2- Logit
3- Fan
4- Palaniswami
5- Min
6- Lee

شین و همکاران[1] (۲۰۰۵)، از شبکه‌های عصبی و ماشین‌های بردار پشتیبان برای پیش‌بینی ورشکستگی شرکت‌ها استفاده کردند و به این نتیجه رسیدند که روش ماشین‌های بردار پشتیبان کارایی بهتری دارد.

پژوهش‌های اخیر در زمینه استفاده از دیگر روش‌های داده‌کاوی مانند k–نزدیک‌ترین همسایه (Park&Han, 2002 and Bian& Mazlack, 2003) و درخت تصمیمنیز نتایج موفقیت آمیزی در زمینه پیش‌بینی درماندگی مالی شرکت‌ها داشته است(Cho, 2010 and Olson, 2012).

علاوه بر این، برخی دیگر نیز از روش‌های ترکیبی داده‌کاوی استفاده کرده‌اند که در ادامه به دو نمونه از آنها اشاره می‌گردد:

هوآ و همکاران[2] (۲۰۰۷) برای پیش‌بینی درماندگی مالی شرکت‌ها از قانون تمایز دو دویی یکپارچه[3] استفاده کردند. ایشان ادعا می‌کنند که با تفسیر و اصلاح خروجی ماشین بردار پشتیبان بر اساس تحلیل رگرسیون لجیستیک، از میزان ریسک تجربی[4] خروجی ماشین بردار پشتیبان کاسته شده است.

لی و همکاران[5] (۲۰۰۹) از استدلال مبتنی بر مورد[6] که یکی از فنون حل مساله است، و نیز روابط ارشدیت[7] استفاده کرده و با استفاده از روش k–نزدیک‌ترین همسایه درماندگی مالی را در شرکت‌های چینی پیش‌بینی کردند.

۳. روش k–نزدیک‌ترین همسایه

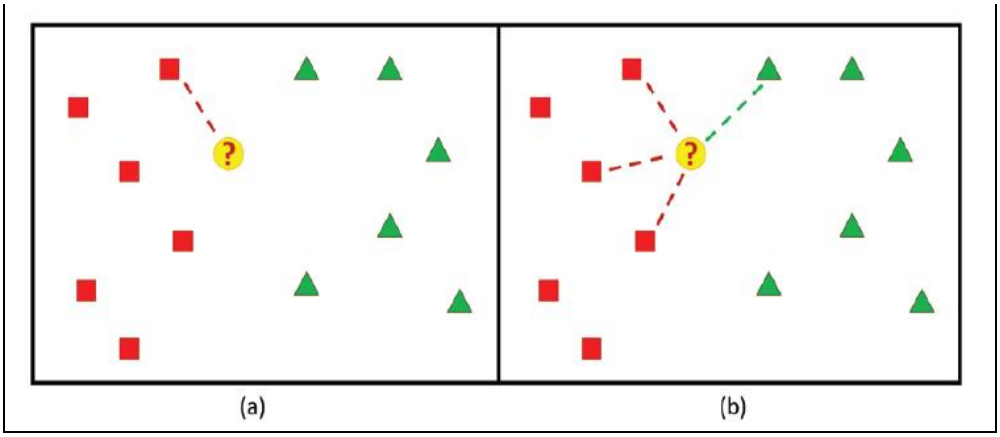
روش k–نزدیک‌ترین همسایه جزو ۱۰ روش برتر در داده‌کاوی است (موچرینو و همکاران[8]، ۲۰۰۹). این روش *وابسته به حافظه*[9] بوده و نیاز به هیچ مدلی برای برازش کردن[1] مدل ندارد. اگر

1- Shin et al.
2- Hua et al.
3- Integrated Binary Discriminant Rule (IBDR)
4- Empirical Risk
5- Li et al.
6- Case-based reasoning (CBR)
7- Outranking relations (OR)
8- Mucherino et al.
9- memory-based

یک نقطه سوال[۲] فرضی (x_0) وجود داشته باشد، k نقطه آموزشی $x_{(r)}, r=1, 2, \ldots, k$ که در نزدیک‌ترین فاصله به (x_0) باشند پیدا کرده و سپس بر اساس رای‌گیری اکثریت[۳] در میان k همسایه دسته‌بندی می‌شود. این روش بسیار ساده است، اما علی‌رغم سادگی در حل مسائل دسته-بندی مانند تشخیص دستخط و تحلیل تصاویر ماهواره‌ای موفق بوده است (هاستی و همکاران[۴]، ۲۰۰۱). به‌وسیله این روش می‌توان یک نمونه ناشناخته را بر اساس دسته‌بندی همسایه‌های آن، دسته‌بندی نمود. در این روش هر نمونه می‌بایست به وسیله نمونه‌های مجاور آن دسته‌بندی شود. بنابراین، اگر دسته‌بندی یک نمونه نامشخص باشد، می‌توان با توجه به دسته‌بندی نزدیک‌ترین نمونه‌های همسایه آن، آنها را پیش‌بینی کرد. با داشتن یک نمونه نامشخص و یک مجموعه آموزشی، تمام فواصل میان آن نمونه و دیگر نمونه‌ها در مجموعه آموزشی قابل محاسبه است. کوچک‌ترین فاصله محاسبه شده با نمونه مورد نظر در مجموعه آموزشی، نزدیک‌ترین فاصله به نمونه ناشناس می‌باشد. بنابراین، می‌توان نمونه ناشناس را بر اساس دسته‌بندی این نزدیک‌ترین همسایه دسته‌بندی نمود.

در حالت کلی، در این روش به هر نمونه دسته‌بندی نشده کلاسی اختصاص داده می‌شود که مشتمل بر k عدد از نزدیک‌ترین همسایه‌های آن می‌باشد. به همین خاطر است که این روش داده کاوی، k–نزدیک‌ترین همسایه نامیده می‌شود. اگر فقط یکی از نمونه‌های مجموعه آموزشی، برای دسته‌بندی مورد استفاده قرار بگیرد، بدان معنی است که ۱–نزدیک‌ترین همسایه اعمال شده است. شکل ۱ قانون تصمیم گیری k–نزدیک‌ترین همسایه را به ازای $k=1$ و $k=4$ برای مجموعه‌ای از نمونه‌ها که به دو دسته تقسیم شده‌اند را نشان می‌دهد.

1- fit
2- Query point
3- Majority Vote
4- Hastie et al.



شکل (۱): (a) قانون تصمیم‌گیری 1-NN: نقطه ؟ به کلاس سمت چپ اختصاص دارد؛ (b) قانون تصمیم‌گیری k-NN، با $k=4$: در اینجا نیز نقطه ؟ به کلاس سمت چپ تعلق دارد.

مأخذ: محاسبات تحقیق.

در شکل ۱-(a) یک نمونه ناشناس تنها با استفاده از یک نمونه مشخص دسته‌بندی شده است. شکل ۱-(b) بیش از یک نمونه مشخص مورد استفاده قرار گرفته است. در این حالت پارامتر k روی ۴ تنظیم شده است، به طوری که ۴ نمونه نزدیک‌تر برای دسته‌بندی نمونه نامشخص مورد استفاده قرار گرفته‌اند. ۳ نمونه از این‌ها به یک کلاس تعلق دارند، در حالی که فقط یکی از نمونه‌ها به کلاس دیگری تعلق دارد. در هر دو حالت، نمونه نامشخص در کلاس سمت چپ دسته‌بندی می‌شود.

تابع فاصله نقش مهمی را در موفقیت دسته‌بندی بازی می‌کند. عامل مهم دیگر، انتخاب مقدار پارامتر k است، از آنجا که این پارامتر تعیین‌کننده تعداد نزدیک‌ترین همسایه‌هاست، لذا مهم‌ترین پارامتر در این روش می‌باشد. اگر k خیلی بزرگ باشد، کلاس‌هایی که تعداد نمونه‌های دسته‌بندی شده آنها خیلی زیاد باشد، کلاس‌هایی را که تعداد نمونه‌های آنها کم باشد را در خود جای داده و در نتیجه نتایج دچار انحراف خواهند شد. از طرف دیگر، اگر k خیلی کوچک باشد، از مزیت استفاده از تعداد نمونه‌های زیاد برای فرایند آموزش بهره کافی برده نخواهد شد (موچرینو و همکاران، ۲۰۰۹). همان‌طور که دومنیکونی و همکاران[1] (۲۰۰۲) نشان داده‌اند، اگر نقاط به طور

1- Domeniconi et al.

یکنواخت توزیع نشوند، محاسبه دقیق مقدار k مشکل خواهد شد. در حالت کلی، مقادیر بزرگتر k نسبت به نویز و اغتشاش ایمن‌ترند و هم‌چنین سبب هموارتر شدن مرز بین کلاس‌های مختلف می‌شوند (Song, 2007).

گفته می‌شود که روش k-NN یک دسته‌بندی کننده *تنبل* است، زیرا در حقیقت این روش از داده‌های موجود در مجموعه داده‌ای برای ایجاد یک دسته‌بندی کننده استفاده نمی‌کند، بلکه به جای آن از داده‌های آموزشی، هر زمان که نیاز به دسته‌بندی باشد، استفاده می‌کند. این امر سبب می‌شود تا این روش نسبت به سایر روش‌ها آسان‌تر باشد، ولی در عوض هزینه محاسبات بیشتر خواهد بود (Mucherino, 2009).

۴. روش تحقیق

انجام هر پژوهش مستلزم تعیین و تعریف هر یک از متغیرهای آن می‌باشد. متغیرها بر اساس نقشی که در پژوهش به عهده دارند به دو دسته تقسیم می شوند:

الف) متغیرهای وابسته: در پژوهش حاضر یک متغیر وابسته وجود دارد که همان وضعیت مالی شرکت‌ها از لحاظ توانمندی مالی است و دارای دو حالت است: سالم یا درمانده.

ب) متغیرهای مستقل: متغیرهای مستقل پژوهش جاری، نسبت‌های مالی می‌باشند. برای انتخاب نسبت‌های مالی مورد استفاده در این پژوهش از مدل چهار متغیره موسوی‌شیری و طبرستانی (۱۳۸۸)، مدل سه متغیره کمیجانی و سعادت‌فر (۱۳۸۵) و نیز مدل شش متغیره نیکبخت و شریفی (۱۳۸۹) استفاده شده است. نسبت‌های مالی مورد استفاده در این سه تحقیق، در جدول ۱ آمده است.

به منظور آزمون جهت استخراج پاسخ‌های لازم به سؤالات اصلی تحقیق، فرضیه‌های تحقیق به شرح زیر مطرح می شود:

فرضیه۱: پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از روش تجزیه و تحلیل نسبت‌ها امکان‌پذیر است.

فرضیه۲: روش k–نزدیک‌ترین همسایه قادر است با دقت خوبی درماندگی مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران را پیش‌بینی کند.

جدول (۱): متغیرهای مورد استفاده در پژوهش

مدل سه‌متغیره	مدل چهارمتغیره	مدل شش‌متغیره
نسبت جاری	دارایی‌های جاری به کل دارایی‌ها	سرمایه در گردش به کل دارایی‌ها
حاشیه سود ناخالص	سرمایه در گردش به کل دارایی‌ها	جریانات نقدی به کل بدهی
		سود خالص به کل دارایی‌ها
نسبت سود خالص به بدهی‌های جاری	سود خالص به کل دارایی‌ها	کل بدهی‌ها به کل دارایی‌ها
	سود ناخالص به فروش	نسبت جاری
		دارایی‌های سریع به بدهی‌های جاری

مأخذ: گردآوری تحقیق

پس از مطالعه و بررسی پژوهش‌های مشابه که در محیط اقتصادی ایران انجام شده‌اند و از میان نسبت‌های مختلف، پنج نسبت مالی که بهترین نتایج را به دست دادند به عنوان متغیرهای مستقل تحقیق انتخاب شدند که عبارتند از:

سرمایه در گردش به کل دارایی‌ها

سود خالص به کل دارایی‌ها

کل بدهی‌ها به کل دارایی‌ها

نسبت جاری

دارایی‌های سریع به بدهی‌های جاری

روش تحقیق مورد استفاده در این پژوهش، از نوع تحقیقات غیرآزمایشگاهی بوده و ملاک طبقه‌بندی شرکت‌ها در دو گروه درمانده و سالم، شمول یا عدم‌شمول ماده ۱۴۱ قانون تجارت می‌باشد. جامعه تحقیق در این پژوهش، شرکت‌های تولیدی پذیرفته در بورس اوراق بهادار تهران طی سال‌های ۱۳۸۶-۱۳۸۴ و روش نمونه‌گیری تصادفی می‌باشد.

در این تحقیق در مجموع از اطلاعات مالی ۱۱۷ شرکت استفاده شده است که به دو دسته تقسیم شده‌اند. دسته نخست، که نمونه آموزشی نام دارد، متشکل از ۷۵ شرکت (۴۳ شرکت سالم و ۳۲ شرکت درمانده) در سال ۱۳۸۶ بوده و دسته دوم که نمونه کنترل نامیده می‌شود، از ۴۲ شرکت (۲۰ شرکت سالم و ۲۲ شرکت درمانده در همان سال) تشکیل شده است.

داده‌ها برای شرکت‌های نمونه، برای یک سال قبل (*t-1*) و دو سال قبل (*t-2*) از وقوع درماندگی مالی بررسی شده و در مجموع از اطلاعات سه سال مالی استفاده شده است. از آن‌جا که سال مبنا (*t*) در این تحقیق سال ۱۳۸۶ می‌باشد، بنابراین در کل داده‌های سه سال مالی ۱۳۸۶-۱۳۸۴ مورد بررسی قرار گرفته است.

همه مدل‌های آماری به‌کار رفته در این تحقیق ابتدا نیاز به آموزش[1] دارند، بنابراین نسبت‌های مالی همه شرکت‌های گروه نخست (گروه آموزشی متشکل از ۷۵ شرکت) در سال مبنا (*t*)، به عنوان داده‌های آموزشی، به مدل‌ها داده شده و سپس از داده‌های سال‌های *t-1* و *t-2* برای آزمون مدل بهره گرفته شد. سرانجام، با مقایسه مقدار پیش‌بینی شده توسط مدل، با مقدار واقعی، دقت مدل‌ها اندازه‌گیری گردید.

با توجه به این‌که در نمونه آموزشی از داده‌های سال‌های گذشته خود شرکت‌ها برای پیش‌بینی استفاده شده است، هدف از ایجاد گروه کنترل، آزمودن قدرت مدل در پیش‌بینی درماندگی مالی هر شرکت جدید خارج از گروه آموزشی می‌باشد.

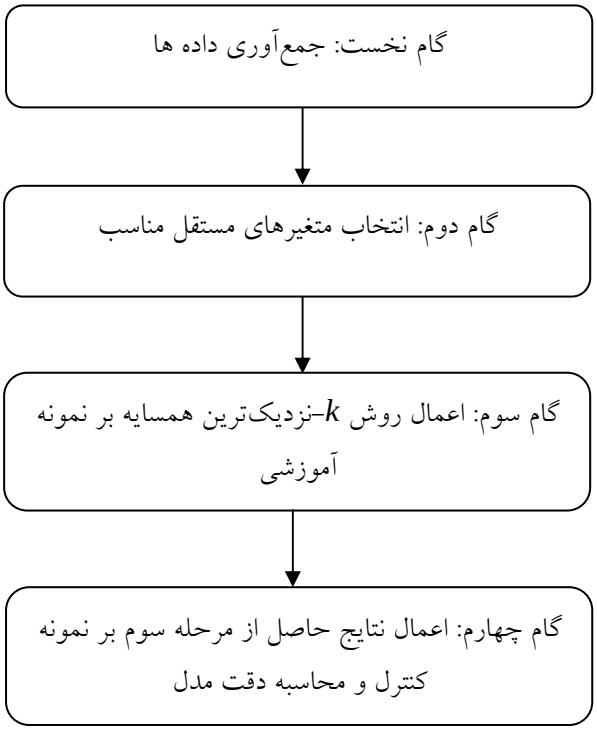
برای گردآوری داده‌ها نیز از صورت‌های مالی موجود در سامانه[2] سازمان بورس و اوراق بهادار[3] و هم‌چنین سامانه مدیریت پژوهش، توسعه و مطالعات اسلامی[4] سازمان بورس و اوراق بهادار استفاده شده است. مراحل انجام تحقیق به طور خلاصه در شکل ۲ نشان داده شده است.

برای ارزیابی قدرت و کارایی مدل‌ها در دسته‌بندی شرکت‌ها، از سه شاخص زیر می‌توان استفاده کرد:

خطای نوع اول (خطای آلفا): هنگامی اتفاق می‌افتد که یک شرکت درمانده در زمره شرکت-های غیر درمانده دسته‌بندی شود.

خطای نوع دوم (خطای بتا): هنگامی اتفاق می‌افتد که یک شرکت غیر درمانده در زمره شرکت‌های درمانده دسته‌بندی شود.

1- Training
2- Website
3- www.tsetmc.com
4- www.rdis.ir



شکل (۲): مراحل انجام تحقیق

دقت کلی

که تعریف این شاخص‌ها در ادامه، در معادلات (۱) تا (۳) آمده است:

خطای نوع اول= تعداد شرکت‌های درمانده که در زمره شرکت‌های غیر درمانده دسته‌بندی شده‌اند/ تعداد شرکت‌هایی که به‌عنوان غیر درمانده دسته‌بندی شده‌اند (۱)

خطای نوع دوم= تعداد شرکت‌های غیر درمانده که در زمره شرکت‌های درمانده دسته‌بندی شده‌اند/ تعداد شرکت‌هایی که به عنوان درمانده دسته‌بندی شده‌اند (۲)

دقت کلی= تعداد شرکت‌هایی که به درستی دسته‌بندی شده‌اند/ تعداد کل شرکت‌ها (۳)

در تحلیل اعتبار[1]، اگر یک مشتری مستعد درماندگی به صورت سالم دسته‌بندی شود، خطای نوع اول به معنی از دست دادن وام و هم‌چنین سود حاصل از آن می‌باشد، اما خطای نوع دوم -که نتیجه دسته‌بندی یک شرکت سالم در زمره درماندگان است- منجر به از دست‌دادن یک مشتری می‌شود. به همین دلیل خطای نوع اول هزینه بیشتری را بر شرکت تحمیل می‌کند (Neves, 2006). بر اساس تحقیقات آلتمن (۱۹۷۷) هزینه‌ای که خطای نوع اول بر شرکت وارد می‌کند، ۳۵ برابر بیشتر از خطای نوع دوم است. بنابراین، می‌توان گفت خطای نوع اول، اهمیت بیشتری نسبت به خطای نوع دوم دارد (Aliakbari, 2009).

در این پژوهش، ۳ معیار معرفی شده برای ارزیابی مدل، مورد استفاده قرار گرفته است، که اولویت اول با میزان دقت کلی بوده و خطای نوع اول و خطای نوع دوم در رتبه‌های بعدی از نظر اهمیت قرار دارند.

۵. یافته‌های تحقیق

در جدول ۲ و جدول ۳ بهترین نتایج حاصل از اعمال مدل k-نزدیک‌ترین همسایه، بر نمونه آموزشی در سال‌های ۱۳۸۴ و ۱۳۸۵ آمده است. در این‌جا، تعداد همسایه‌ها و روش‌های اندازه-گیری فاصله مهم‌ترین عوامل تعیین‌کننده میزان دقت مدل می‌باشند. همان‌گونه که مشاهده می-شود، اگر تعداد همسایه‌ها ۷ یا ۸ و روش اندازه‌گیری فاصله متناظر با آن، به‌ترتیب Chebychev، Euclidean یا Euclidean squared باشد، بهترین نتیجه با دقت ۹۰/۶۷٪، برای سال t-1 (۱۳۸۵) به دست می‌آید. برای سال t-2 (۱۳۸۴) نیز بهترین نتیجه هنگامی حاصل می‌گردد که تعداد همسایه‌ها ۸ و روش اندازه‌گیری Chebychev باشد. دقت بهترین نتیجه حاصله در این سال برابر با ۸۴٪ است.

پس از اعمال نتایج حاصل از نمونه آموزشی (تعداد همسایه‌ها و تابع اندازه‌گیری فاصله متناظر با آن) بر نمونه کنترلی، میزان دقت کلی در یک سال و دو سال قبل از سال مبنا به‌ترتیب برابر با ۸۵/۷۱٪ و ۹۵/۲۴٪ به دست آمد که این نتایج در جدول ۴ آمده است.

1- Credit Analysis

جدول (۲): نتایج حاصل از اعمال k-نزدیک‌ترین همسایه بر نمونه آموزشی در سال ۱۳۸۵

تعداد همسایه‌ها	۷	۷	۷	۸
روش اندازه‌گیری فاصله	Euclidean	Euclidean squared	Cityblock (Manhattan)	Chebychev
میزان دقت مدل	٪۹۰/۶۷	٪۹۰/۶۷	٪۸۹/۳۳	٪۹۰/۶۷
خطای کلی	٪۹/۳۳	٪۹/۳۳	٪۱۰/۶۷	٪۹/۳۳
خطای نوع اول	٪۱۳/۴۶	٪۱۳/۴۶	٪۱۵/۰۹	٪۱۳/۴۶
خطای نوع دوم	٪۰	٪۰	٪۰	٪۰

مأخذ: محاسبات تحقیق

جدول (۳): نتایج حاصل از اعمال k-نزدیک‌ترین همسایه بر نمونه آموزشی در سال ۱۳۸۴

تعداد همسایه‌ها	۷	۷	۷	۸
روش اندازه‌گیری فاصله	Euclidean	Euclidean squared	Cityblock (Manhattan)	Chebychev
میزان دقت مدل	٪۸۰	٪۸۰	٪۸۱/۳۳	٪۸۴
خطای کلی	٪۲۰	٪۲۰	٪۱۸/۶۷	٪۱۶
خطای نوع اول	٪۲۵/۴۲	٪۲۵/۴۲	٪۲۴/۱۴	٪۲۰/۳۷
خطای نوع دوم	٪۰	٪۰	٪۰	٪۴/۷۶

مأخذ: محاسبات تحقیق

جدول (۴): نتایج حاصل از اعمال k-نزدیک‌ترین همسایه بر نمونه کنترلی در سال‌های ۱۳۸۴ و ۱۳۸۵

سال	۱۳۸۵				۱۳۸۴	
روش اندازه-گیری فاصله	Euclidean	Euclidean squared	Cityblock (Manhattan)	Chebychev	Euclidean	Euclidean squared
میزان دقت مدل	٪۸۵/۷۱				٪۹۵/۲۴	
خطای کلی	٪۱۴/۲۹				٪۴/۷۶	
خطای نوع اول	٪۲۴				٪۰	
خطای نوع دوم	٪۰				٪۱۰	

مأخذ: محاسبات تحقیق

از آن جا که نمونه با اهمیت در پژوهش جاری نمونه کنترلی است، با توجه به نتایج جدول اخیر پیش‌بینی درماندگی مالی شرکت‌ها در محیط اقتصادی ایران امکان پذیر بوده و روش k- نزدیک‌ترین همسایه قادر است با دقت خوبی (۸۵/۷۱٪ در یک سال و ۹۵/۲۴٪ در دو سال قبل از سال مبنا) درماندگی مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران را پیش‌بینی کند. لذا هر دو فرضیه تحقیق تایید می‌گردند.

۶. جمع بندی و نتیجه‌گیری

با توجه به هزینه‌های بالای ورشکستگی و همچنین رابطه بین درماندگی مالی شرکت‌ها و ورشکستگی آنها، می‌توان با پیش‌بینی به موقع درماندگی مالی از ورشکستگی آنها جلوگیری کرد.

از آن‌جا که در سال‌های اخیر و با پیشرفت‌های صورت گرفته در علوم کامپیوتر، به جای استفاده از مدل‌های سنتی آماری، مدل‌های داده‌کاوی با توجه دقت بالا و هم‌چنین خواص غیرخطی و ناپارامتری آنها بیشتر مورد توجه قرار گرفته است، لذا در پژوهش جاری از روش k- نزدیک‌ترین همسایه برای پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده در بوس اوراق بهادار تهران استفاده شد.

پس از استخراج صورت‌های مالی شرکت‌ها، مطالعه و بررسی تحقیقات مشابه و استخراج صورت‌های مالی مورد استفاده در آنها و هم‌چنین آزمودن ترکیبات مختلف، در نهایت پنج نسبت مالی شامل سرمایه در گردش به کل دارایی‌ها، سود خالص به کل دارایی‌ها، کل بدهی‌ها به کل دارایی‌ها، نسبت جاری و دارایی‌های سریع به بدهی‌های جاری بهترین نتایج را به دست دادند.

در مرحله بعد، داده‌های سال مبنا به عنوان داده‌های آموزشی به مدل‌ها داده شد و با استفاده از آن و پس از به دست آوردن بهترین نتایج در نمونه آموزشی، تعداد همسایه‌ها و تابع اندازگیری فاصله به دست آمد. سپس، از این مشخصه‌ها در نمونه کنترلی استفاده گردید، تا بدین‌وسیله قدرت مدل سنجیده شود. نتایج به دست آمده نشان می‌دهد که روش داده‌کاوی مورد استفاده در این تحقیق با دقت خوبی قادر به دسته‌بندی شرکت‌های سالم و درمانده می‌باشد.

نتایج تحقیق اخیر نشان می‌دهد که پیش‌بینی پدیده درماندگی مالی در محیط اقتصادی ایران امکان‌پذیر است. از آن‌جا که این پیش‌بینی با استفاده از اطلاعات مالی موجود در صورت‌های

مالی شرکت‌ها انجام شده است، این امر می‌تواند مهر تاییدی بر وجود اطلاعات کافی در صورت-های مالی شرکت‌ها، برای پیش‌بینی وضعیت مالی آتی آنها باشد.

هم‌چنین، نتایج حاصل از این پژوهش حاکی از آن است که نسبت‌های مالی می‌توانند پیش-بینی‌کننده خوبی برای درماندگی مالی شرکت‌ها باشند. به علاوه، با توجه به توان بالای مدل ارائه شده، بانک‌ها و موسسات رتبه‌بندی اعتباری می‌توانند از مدل مورد استفاده در این پژوهش برای رتبه‌بندی اعتباری شرکت‌ها استفاده نمایند.

در مجموع باید گفت که با استفاده از نتایج این پژوهش به‌عنوان اولین گام، می‌توان با تشخیص به موقع، از مبتلا شدن شرکت‌ها به درماندگی مالی و هم‌چنین پیامدهای آن جلوگیری کرد.

*این تحقیق با حمایت دانشگاه پیام نور انجام شده است.

منابع

[1] Aliakbari, S. (2009) "Prediction of Corporate Bankruptcy for the UK Firms in Manufacturing Industry", London: Brunel University, MSc dissertation in Finance and Investment.

[2] Altman, E.I. (1968) "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", The Journal of Finance, 23(4): 589–609.

[3] Altman, E. I., Haldeman, R. G. & Narayanan, P. (1977) "Zeta analysis: A new model to identify bankruptcy risk of corporations" Journal of Banking and Finance, 1: 29–54.

[4] Beaver, W.H. (1966) "Financial Ratios as Predictors of Failure", Journal of Accounting Research, 4: 71-111.

[5] Bian, H. & Mazlack, L. (2003) "Fuzzy-rough nearest-neighbor classification approach", Proceeding of 22nd International Conference of the North American Fuzzy Information Processing Society (NAFIPS 2003), IEEE, Los Alamitos, 500-505.

[6] Cho, S., Hong, H. & Ha, B.C. (2010)"A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance: For bankruptcy prediction", Expert Systems with Applications, 37(4): 3482-3488.

[7] Domeniconi, C., Peng, J. & Gunopulos, D. (2002) "Locally adaptive metric nearest-neighbor classification", IEEE Transactions on Pattern Analysis and Machine Intelligence. 24(9): 1281–1285.

[8] Elmer, P.J., & Borowski, D.M. (1988) "An expert system approach to financial analysis: The case of S&L bankruptcy", Financial Management, 17(3): 66–76.

[9] Fan, A. & Palaniswami, M. (2000) "A new approach to corporate loan default prediction from financial statements", Proceedings of the computational finance/forecasting financial markets conference (CF/FFM-2000), London, UK.
[10] Hastie, T., Tibshirani, R. & Friedman, J. (2001) "The elements of statistical learning: data mining, inference, and prediction", New York: Springer.
[11] Jensen, H.L. (1992) "Using neural networks for credit scoring", Managerial Finance, 18: 15–26.
[12] Hua, Z., Wang, Y., Xu, X., Zhang, B. & Liang, L. (2007) "Predicting corporate financial distress based on integration of support vector machine and logistic regression", Expert Systems with Applications, 33(2): 434-440.
[13] Komijani, A. & Saadatfar, J. (2007) "Application of Artificial Neural Network models in predicting economic bankruptcy of the registered corporations in stock market", 3(6): 11-44 (in Persian).
[14] Li, H., Sun, J. & Sun, B.L. (2009) "Financial distress prediction based on OR-CBR in the principle of k-nearest neighbors" Expert Systems with Applications, 36(1): 643-659.
[15] Malhotra, R. & Malhotra, D.K. (2002) "Differentiating between good credits and bad credits using neuro-fuzzy systems", European Journal of Operational Research, 136(2): 190–211.
[17] Markham, I.S., & Ragsdale, C.T. (1995) "Combining neural networks and statistical predictions to solve the classification problem in discriminant analysis", Decision Sciences, 26(2): 229–242.
[18] Min, J.H. & Lee, Y.C. (2005) "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters", Expert Systems with Applications, 28(4): 603-614.
[19] Mousavi Shiri, M. & Tabarestani, M.R. (2010) "Predicting financial distress by using DEA", Accounting Research, 2: 158-187 (in Persian).
[20] Mucherino, A., Papajorgji, P.J. & Pardalos, P.M. (2009) "Data mining in agriculture", New York: Springer.
[21] Neves, J. C., Vieira, A. (2006) "Improving bankruptcy prediction with hidden layer learning vector quantization" European Accounting Review, 15(2): 253–271.
[22] Nikbakht, M.R., & Sharifi, M. (2010) "Predicting financial bankruptcy of Tehran Stock Exchange Firms by using Artificial Neural Networks", Industrial Management magazine, 2(4): 163-180 (in Persian).
[23] Ohlson, J.A. (1980). "Financial ratios and the probabilistic prediction of bankruptcy" Journal of Accounting Research, 18(1): 109–131.
[24] Olson, D. L., Delen, D. & Meng, Y. (2012) "Comparative analysis of data mining methods for bankruptcy prediction", Decision Support Systems, 52(2): 464-473.
[25] Park, C.S. & Han, I. (2002) "A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction", Expert Systems with Applications, 23(3): 255–264.

[26] Rasoolzadeh, M. (2002) "Application of Altman's model in determining corporate bankruptcy, Tadbir Magazine, 120: 105-107 (in Persian).
[27] Shin, K., Lee, T.S. & Kim, H. (2005) "An application of support vector machines in bankruptcy prediction model", Expert Systems with Applications, 28(1): 127-135.
[28] Song Y., Huang, J., Zhou, D., Zha, H. & Giles, C. (2007) "IKNN: Informative K-Nearest Neighbor", Pattern Classification, Springer-Verlag Berlin Heidelberg, 248–264.
[29] Srinivasan, V. & Ruparel, B. (1990) "CGX: An expert support system for credit granting", European Journal of Operational Research, 45: 293–308.
[30] Sun, J., Li, H. (2008) "Listed companies' financial distress prediction based on weighted majority voting combination of multiple classifiers", Expert Systems with Applications 35: 818–827.
[31] Sun, J., Li, H. (2009) "Financial distress prediction based on serial combination of multiple classifiers", Expert Systems with Applications, 36: 8659–8666.
[32] Zhang, G., Hu, M.Y., Patuwo, B. E. & Indro, D.C. (1999) "Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis", European Journal of Operational Research 116(1): 16-32.