

کاربرد روش‌های رده‌بندی در پیش‌بینی ریزش مشتری

سیدعلیرضا مهدوی تالارپشتی - داود شاهسونی

گروه آمار، دانشکده ریاضی، دانشگاه صنعتی شاهرود

چکیده: در این مقاله، موضوع پیش‌بینی ریزش مشتری^۱ در صنعت تلفن همراه توسط سه مدل شبکه‌های عصبی، رگرسیون لجستیک و درخت تصمیم مورد مطالعه قرار گرفته و کارایی این مدل‌ها توسط چندین معیار با یکدیگر مقایسه شده است. نتایج نشان داد که مدل‌های شبکه‌های عصبی و درخت تصمیم دارای بیشترین دقت برای داده‌های آزمون هستند. با این وجود، از آنجا که شناسایی مشتریانی که ریزش آنها محتمل است از جمله اهداف مدیریتی می‌باشد، مدل رگرسیون لجستیک به عنوان یک گزینه مناسب پیشنهاد می‌گردد.

واژه‌های کلیدی: ریزش مشتری، شبکه‌های عصبی، درخت تصمیم، رگرسیون لجستیک

۱ مقدمه

در سال‌های اخیر تعداد شرکت‌هایی که خدمات نسبتاً مشابهی را به جامعه ارائه می‌دهند موجب شده است تا مشتریان بر اساس نیاز خود بتوانند شرکت مورد نظر را انتخاب و از خدمات آن بهره‌مند گردند. این گوناگونی به وضوح در صناعی از قبیل بیمه، مخابرات، موسسات ارائه‌دهنده خدمات اینترنت و شبکه‌های تلویزیون کابلی دیده می‌شود و سبب شده است تا موضوع نگاه‌داشتن مشتریان فعلی و عدم جذب آنها توسط غیر، از جمله مهم‌ترین استراتژی‌های مدیریتی تلقی شود. موضوع قطع ارتباط مشتری با سرویس‌دهنده فعلی، موسوم به ریزش مشتری است. از دست دادن یک مشتری سودآور به معنی جذب او توسط سرویس‌دهنده‌ی رقیب بوده و هزینه‌ی جذب مشتری جدید به مراتب بیشتر از هزینه‌ی نگهداری آن است (Mozer, 2002). از دیدگاه مدیریت ریسک و اقتصاد، تشخیص مشتریانی که ریزش آنها مخاطره‌زایی دارد، بسیار مهم و قابل توجه است. اطلاعات موجود در سوابق مشتریان، اعم از مشتریان وفادار و ریزش‌شده، مبنایی برای پیش‌گویی رفتار آینده مشتریان است. اگر بتوان بر اساس اطلاعات

^۱ Customer Churn

مربوط به ویژگی‌های مشتریان، احتمال ریزش یا عدم آن را پیش‌بینی کرد، می‌توان با انجام فعالیت‌های بازدارنده، ریزش آنها را به حداقل رسانید. به منظور پیش‌بینی ریزش مشتری، روش‌های گوناگون آماری و یادگیری ماشین^۲ برای رده‌بندی وجود دارد که از جمله مهمترین آنها می‌توان به رگرسیون لجستیک، شبکه‌های عصبی و درخت تصمیم اشاره کرد. در این مطالعه مدل‌های مذکور را معرفی نموده و دقت و کارایی آنها را در پیش‌بینی ریزش مشتری مورد بررسی قرار می‌دهیم.

۲ مروری کوتاه بر مطالعات قبلی

پیش‌بینی ریزش مشتری برای سازمان‌ها و کمپانی‌هایی که حیات آنها به وجود مشتری- نشان وابسته است، بسیار حائز اهمیت است. در سال‌های اخیر مطالعات بسیاری در خصوص ریزش مشتری در صنایعی چون بیمه، بانکداری، مخابرات و فروشگاه‌های بزرگ زنجیره‌ای انجام گرفته است. در هر یک از این مطالعات از مدل‌های مختلف آماری و یادگیری ماشین استفاده شده و کارایی آنها مورد ارزیابی قرار گرفته است. رشید (Rashid, ۲۰۰۹) با استفاده از تکنیک آموزش پویا، سعی به بهبود دقت شبکه عصبی بازگشتی نمود و سپس با استفاده از آن، ریزش مشتری را در شرکت‌های ارتباطاتی پیش‌بینی کرده است. شارما و همکاران (Sharma, ۲۰۱۱) با استفاده از مدل شبکه عصبی، ریزش مشتریان را برای اپراتورهای تلفن همراه با دقتی بیش از ۹۲ درصد، پیش‌بینی کردند. کرامتی و همکاران (Keramati, ۲۰۱۱) نیز از رگرسیون لجستیک به منظور پیش‌بینی ریزش مشتری در اپراتور تلفن همراه در ایران استفاده کرده و به این نتیجه رسیده‌اند که متغیرهای مربوط به برخی ویژگی‌های جمعیتی مشتریان، در تصمیم-گیری آنها برای ترک کردن یا ادامه رابطه با کمپانی بسیار موثر است.

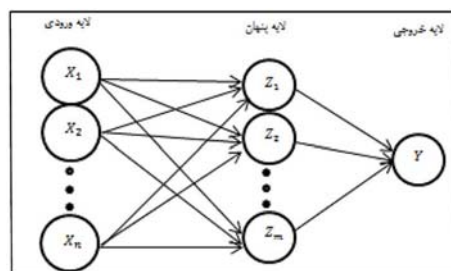
۳ معرفی مدل‌های مورد مطالعه

در این بخش به معرفی مدل‌های شبکه عصبی مصنوعی، رگرسیون لجستیک و درخت تصمیم که در این تحقیق مورد استفاده واقع شده است می‌پردازیم. در همه این مدل‌ها متغیرهای توضیحی با $X = (X_1, X_2, \dots, X_n)$ و متغیر پاسخ کیفی با Y نشان داده می‌شوند و متغیر پاسخ مقادیر ۱ و ۰ را به ترتیب به ازای ریزش و عدم ریزش مشتری اختیار می‌کند.

^۲ Machine learning

۱.۳ شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی که از این پس به اختصار آن را شبکه‌های عصبی می‌نامیم، رده‌ای از روش‌های یادگیری است که در زمینه‌های آماری و هوش مصنوعی رشد و توسعه یافته است و دارای قابلیت رگرسیونی و رده‌بندی می‌باشد. در مدل شبکه عصبی، متغیرهای توضیحی و متغیر پاسخ می‌توانند از نوع کمی یا کیفی باشند و همچنین متغیر پاسخ، رابطه‌ای غیرخطی و غیرمستقیم با متغیرهای توضیحی دارد. چارچوب شبکه‌های عصبی از چند لایه به نام‌های لایه ورودی، لایه(های) پنهان و لایه خروجی تشکیل شده است. شکل ۱، ساختار شماتیکی یک شبکه را با یک لایه پنهان نشان می‌دهد که در آن متغیرهای توضیحی و متغیر پاسخ به ترتیب در لایه ورودی و لایه خروجی واقع شده و اجزای لایه پنهان، Z_i ها، نورون یا گره^۳ نامیده می‌شوند.



شکل ۱: شمایی از ساختار شبکه عصبی

شبکه‌های عصبی را از لحاظ نحوه‌ی اتصال گره‌ها به دو نوع تقسیم می‌کنند.

- ۱) شبکه‌های عصبی پیشرو: که در آن گره‌های هر لایه فقط به گره‌های لایه بعدی متصل می‌شوند (شکل ۱).
- ۲) شبکه‌های عصبی پسرو: که در آن گره‌های هر لایه به گره‌های لایه‌های بعدی و یا به خودشان متصل می‌شوند.

در این مطالعه از شبکه‌های عصبی پیشرو استفاده شده است. یکی از متداول‌ترین سبک‌های معماری شبکه‌های عصبی که در آن الگوی ارتباطی بین لایه‌ها مشخص می‌شود، پرسپترون چندلایه^۴ است. در این ساختار، هر نورون، تابعی از ترکیب خطی متغیرهای

^۳ Node

^۴ Multilayer Perceptron

توضیحی، و همچنین متغیر پاسخ نیز تابعی از ترکیب خطی نوروها است. یعنی:

$$Z_i = \sigma \left(\alpha_{\circ,i} + \sum_{j=1}^n \beta_{j,i} X_j \right); \quad i = 1, 2, \dots, m, \quad y = f \left(\gamma + \sum_{i=1}^m \theta_i Z_i \right) \quad (1)$$

که در آن σ تابع فعال‌سازی نامیده شده و نمودار این تابع غیرخطی، S-شکل^۵ است که از جمله آنها می‌توان به تانژانت هایپربولیک، آرک تانژانت، سینوس و گوسین اشاره کرد. همچنین تابع f ، خطی یا غیرخطی بوده و در صورت غیرخطی بودن می‌تواند همانند σ دارای یکی از فرم‌های مذکور باشد. در برآورد پارامترهای مدل ابتدا مقادیر اولیه این پارامترها به تصادف انتخاب شده و سپس طی یک فرآیند یادگیری، این مقادیر دائما بروز می‌شوند تا زمانی که میزان خطای رده‌بندی به حالت نسبتا پایداری برسد (Hastie، و همکاران، ۲۰۰۹). نهایتا شبکه عصبی با محاسبه احتمال شرطی قرار گرفتن مشاهده‌ی جدید x_0 در کلاس \circ یا ۱ که به آن احتمال پسین گویند، ریزش مشتری را پیش‌بینی می‌کند، به این‌صورت که اگر $P(Y = ۱ | x_0) > 0.5$ ، آنگاه مشاهده‌ی جدید x_0 به کلاس ۱ و در غیر این‌صورت به کلاس \circ تخصیص می‌یابد.

۲.۳ رگرسیون لجستیک

یکی دیگر از ابزارهای بسیار مفید آماری برای رده‌بندی متغیر پاسخ، رگرسیون لجستیک است. اگر متغیر پاسخ دارای دو سطح باشد، به آن رگرسیون لجستیک دودوئی گفته می‌شود و در صورت داشتن بیشتر از دو سطح، بسته به اینکه ترتیب بین سطوح آن مهم باشد یا نباشد، آن را به ترتیب رگرسیون لجستیک ترتیبی یا اسمی نامند. از آنجائیکه متغیر پاسخ در پیش‌بینی ریزش مشتری بصورت ریزش (۱) و یا عدم ریزش (۰) است، لذا در این مطالعه از رگرسیون لجستیک دودوئی استفاده می‌کنیم. برای مشاهده $x = (x_1, x_2, \dots, x_n)$ ، مدل پیش‌بینی کلاس متغیر پاسخ به صورت زیر است:

$$P(x) = E(Y|X = x) = P(Y = ۱|X = x) = \frac{e^{-M}}{1 + e^{-M}} \quad (2)$$

$$Q(x) = P(Y = 0|X = x) = -P(x) = \frac{1}{1 + e^{-M}} \quad (3)$$

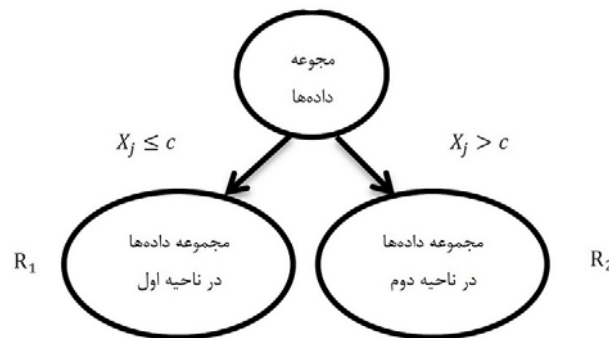
که در آن $M = \alpha + \beta x$ و α و $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ ضرایب ثابت متغیرهای توضیحی هستند که توسط مینیم‌سازی تابع درستمایی مربوطه بدست می‌آیند. لازم به ذکر است که $(Y|X)$ دارای توزیع برنولی با پارامتر $P(x)$ است. پس از برآورد پارامترها، اگر

^۵ Sigmoid Activation Function

$\hat{P}(x_0) > \hat{Q}(x_0)$ ، آنگاه کلاس متغیر پاسخ مربوط به این مشاهده، ۱ پیش‌بینی می‌شود و در غیراینصورت این مشاهده را به کلاس ۰ نسبت می‌دهیم.

۳.۳ درخت تصمیم

درخت تصمیم یکی از تکنیک‌های مرسوم داده‌کاوی است که دارای ساختار درخت شکل بوده و قادر به تولید قانون‌هایی بر اساس مقادیر صفت‌های توضیحی، برای پیش‌بینی کلاس متغیر پاسخ می‌باشد. در درخت تصمیم، فضای متغیرهای توضیحی به ابرمکعب‌های n بعدی مجاور هم‌افراز می‌شوند که هر کدام از این ناحیه‌ها شامل بخشی از مجموعه داده‌ها هستند. سپس در هر ناحیه، کلاسی که دارای فراوانی بیشتری است بعنوان کلاس غالب برای پیش‌بینی متغیر پاسخ در نظر گرفته می‌شود. یعنی اگر مشاهده جدید در یکی از ناحیه‌ها قرار گیرد، آنگاه برای متغیر پاسخ متناظر این مشاهده، کلاسی پیش‌بینی می‌شود که دارای بیشترین فراوانی در آن ناحیه است. نحوه‌ی افراز فضای متغیرهای توضیحی، سلسله‌مراتبی و بصورت دودوئی است بطوریکه در هر مرحله مجموعه داده‌ها به دو قسمت تقسیم می‌شود. این افراز در امتداد یکی از متغیرهای توضیحی و به‌ازای مقدار خاصی از آن متغیر صورت می‌گیرد. فرض کنید فضای متغیرها در راستای متغیر X_j و به‌ازای مشاهده‌ی $x_j = c$ بصورت شکل ۲ افراز شده باشد.



شکل ۲: روند افراز در درخت تصمیم

که در آن ناحیه اصلی شامل کل داده‌ها و R_1 و R_2 بترتیب نواحی چپ و راست ایجاد شده هستند. اگر p_t را فراوانی نسبی کلاس ۱ در ناحیه t در نظر بگیریم، آنگاه شاخص جینی که معرف میزان ناخالصی ناحیه t است، به صورت رابطه (۴) تعریف می‌شود.

$$i(t) = p_t(1 - p_t) \quad (۴)$$

اگر مطابق شکل ۲، فضای متغیرها در راستای جهت متغیر X_j و از نقطه $x_j = c$ به دو قسمت تقسیم شود، تغییر در ناخالصی را می توان به صورت رابطه (۵) تعریف نمود.

$$\Delta i(t) = i(t_R) - E[i(t_c)] = i(t_R) - P_{R_1} i(t_{R_1}) - P_{R_2} i(t_{R_2}) \quad (5)$$

که در آن P_{R_1} و P_{R_2} نسبتی از کل مشاهدات ناحیه R است که به ترتیب در R_1 و R_2 واقع شده اند. با تکرار فرآیند فوق برای تمامی متغیرها و به ازای مقادیر مختلف آنها، مقدار $\Delta i(t)$ برای همه افرازه های ممکن محاسبه شده و نهایتاً افزایی که مقدار بیشینه $\Delta i(t)$ را ایجاد کند به عنوان افراز بهینه در نظر گرفته خواهد شد. بعبارت دیگر باید عمل ماکزیم سازی در فرمول (۶) برای انتخاب راستای تقسیم فضا و نقطه تقسیم انجام شود.

$$\operatorname{argmax}_{X_j \leq x_j, k, j=1, 2, \dots, n, k=1, 2, \dots, m} \Delta i(t) \quad (6)$$

سرانجام، ناحیه های جدید ایجاد شده نیز به همین ترتیب به نواحی کوچکتر تقسیم می شوند تا جائیکه هر درخت دارای تعداد مشاهدات از پیش تعیین شده ای شود.

۴ معیارهای ارزیابی مدل

برای سنجش دقت و کارایی مدل های رده بندی داده ها، از معیارهای مختلفی استفاده می شود که در اینجا بعضی از این معیارها را معرفی می کنیم. فرض کنید مشاهدات دارای کلاس ۱ بعنوان مشاهده ی مثبت (P) و مشاهدات با کلاس ۰ بعنوان مشاهده ی منفی (N) در نظر گرفته شده اند. در جدول ۱ تعداد این مشاهدات بر اساس رده پیش بینی آنها نمادگذاری شده اند که در آن TP تعداد مشاهدات مثبتی است که مثبت پیش بینی شده اند، FP تعداد مشاهدات منفی است که مثبت پیش بینی شده اند، FN تعداد مشاهدات مثبتی است که منفی پیش بینی شده اند و TN تعداد مشاهدات منفی است که منفی پیش بینی شده اند.

جدول ۱: ماتریس حالات پیش بینی برای مسائل رده بندی با دو کلاس

	مشاهده منفی مشاهده مثبت	
پیش بینی مثبت	TP	FP
پیش بینی منفی	FN	TN

برخی از معیارهای ارزیابی مدل های رده بندی عبارتند از:

$$Recall = TPrate = \frac{TP}{TP + FN} \quad (7)$$

$$Precision = \frac{TP}{TP + FP} \quad (۸)$$

$$F - Measure = \sqrt{Precision * Recall} \quad (۹)$$

$$FNrate = \frac{FN}{TP + FN} \quad (۱۰)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (۱۱)$$

که در آن Recall نسبتی از کل مشاهدات مثبت است که درست طبقه‌بندی شده‌اند، Precision نسبتی از پیش‌بینی‌های مثبت است که واقعا مثبت بوده‌اند، FNrate نسبتی از کل مشاهدات مثبت است که اشتباه طبقه‌بندی شده‌اند و میزان دقت یا Accuracy معرف نسبت مشاهداتی است که درست پیش‌بینی شده‌اند. همچنین اندازه F، میانگین هندسی Recall و Precision است که کارایی Recall را نسبت به Precision کنترل می‌کند. مقادیر بزرگ معیارهای Precision، Recall، F، Accuracy و مقادیر کوچک معیار FNrate مبین مناسب‌تر بودن مدل و عملکرد بهتر آن است.

۵ مجموعه داده‌ها

در این مطالعه از مجموعه داده‌های ریزش مشتری موجود در پایگاه داده دانشگاه ایروین^۶ کالیفرنیا (Blake, ۱۹۹۸) استفاده شده است. این مجموعه داده که مربوط به مشترکان شبکه‌ی تلفن همراه است، شامل ۳۳۳۳ مشاهده و ۱۴ متغیر است که ۱۳ متغیر آن، توضیحی و یک متغیر، بعنوان متغیر پاسخ در نظر گرفته شده است. هیچ‌یک از متغیرها دارای مقادیر گمشده نیستند و مشاهدات بطور کامل ثبت شده‌اند. متغیر پاسخ، مقادیر ۱ و ۰ را متناسب با ریزش و عدم ریزش مشتریان اختیار می‌کند. ابتدا داده‌ها را بطور تصادفی به دو دسته تقسیم کرده و ۷۰ درصد آنها را بعنوان داده‌های مدل‌ساز و ۳۰ درصد مابقی را بعنوان داده‌های آزمون در نظر می‌گیریم. تعداد داده‌های مدل‌ساز ۲۳۳۳ مشاهده است که در ۸۵/۲ درصد، متغیر پاسخ مقدار ۰ و در ۱۴/۸ درصد مقدار ۱ را اختیار کرده است که می‌توان گفت نسبت ریزش مشتری به عدم ریزش، ۱ به ۶ است. هنگامیکه یکی از سطوح متغیر پاسخ نسبت به سطوح دیگر دارای فراوانی بسیار بیشتری باشد، اصطلاحاً حالت عدم تعادل یا عدم بالانس^۷ پیش می‌آید. در بروز چنین مواردی الگوریتم‌های یادگیری و روش‌های آماری، عملکرد خوبی نداشته و دقت آنها کاهش می‌یابد و لذا نتایج قابل اعتمادی ارائه نمی‌دهند. برای حل این مشکل،

^۶ Irvine

^۷ Unbalanced situation

روش‌های مختلفی وجود دارد که یکی از آنها روش بیش‌نمونه‌گیری^۸ است. در این روش ساده با تکرار مشاهدات کلاسی که فراوانی کمتری دارند، فراوانی مشاهدات تعدیل می‌شود (Garcia, ۲۰۰۷). تکنیک‌های مختلفی برای انجام بیش‌نمونه‌گیری وجود دارد که یکی از مهمترین آنها، الگوریتم اسموت است که توسط چاولا پیشنهاد شده است (Chawla, ۲۰۰۲). در این الگوریتم بمنظور افزایش فراوانی کلاسی که دارای فراوانی کمتر است، از مقادیر متغیرهای توضیحی مربوط به آن کلاس استفاده شده و با ادغام آنها مشاهدات جدید ایجاد می‌گردد. در این مقاله، با استفاده از الگوریتم اسموت داده‌های مورد مطالعه بالانس شده‌اند که در نتیجه تعداد داده‌های مدل‌ساز به ۳۹۸۴ رسیده و نسبت مشاهدات مثبت به مشاهدات منفی تقریباً برابر ۱ شده‌اند.

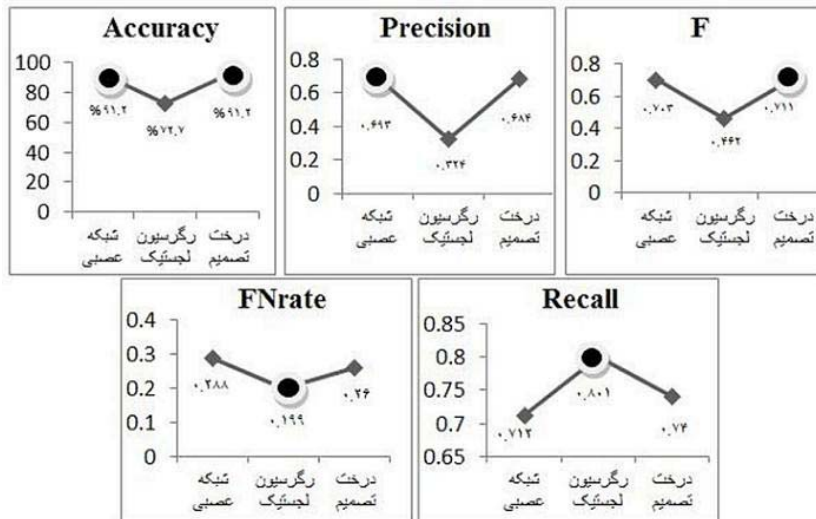
۶ نتایج

در این بخش به ذکر نتایج حاصل از اعمال مدل‌های معرفی شده بر داده‌های ریزش مشتریان شبکه تلفن همراه می‌پردازیم. ابتدا کلیه مدل‌ها را بر داده‌های مدل‌ساز تعدیل شده برازش داده و سپس عملکرد آنها را توسط داده‌های آزمون و با استفاده از معیارهای سنجش دقت که در بخش (۴) معرفی شده‌اند، مورد ارزیابی قرار دادیم. لازم به ذکر است که برای مدل شبکه عصبی، با بررسی‌های انجام شده به این نتیجه رسیدیم که مدلی با تابع فعال‌سازی تانژانت هایپربولیک و یک لایه پنهان با ۳ نورون، نسبت به سایر شبکه‌ها از کارایی بیشتری برخوردار است. نتایج معیارهای سنجش دقت مدل مربوط به داده‌های آزمون در شکل ۳ نمایش داده شده است. با توجه به شکل ۳، می‌توان گفت از لحاظ معیارهای FNrate و Recall بهترین مدل، رگرسیون لجستیک است. همچنین شبکه‌های عصبی بیشترین مقدار Precision را داراست. درخت تصمیم نیز از لحاظ معیار F دارای بیشترین دقت است و نهایتاً بهترین مدل‌ها از لحاظ معیار Accuracy، مدل‌های شبکه عصبی و درخت تصمیم هستند. با توجه به نتایج بدست آمده، هر یک از مدل‌ها دارای نقاط قوت و ضعف مختلفی از لحاظ معیارهای ارزیابی مدل هستند و بسته به نوع مسئله و اهمیت هر یک از معیارهای ارزیابی، می‌توان یکی از مدل‌های معرفی شده را انتخاب نمود.

۷ بحث و نتیجه‌گیری

در این مقاله پیش‌بینی ریزش مشتری در صنعت تلفن همراه، توسط چندین روش آماری و یادگیری ماشین مورد مطالعه قرار گرفت. کارایی هر یک از روش‌های استفاده شده

^۸ Over sampling



شکل ۳: معیارهای سنجش دقت مدل‌ها

براساس معیارهای سنجش دقت، متفاوت است. رگرسیون لجستیک از لحاظ معیارهای Precision و FNrate، بهتر از سایر مدل‌ها عمل کرده است. از لحاظ معیار F و شبکه‌های عصبی و درخت تصمیم بترتیب دارای بیشترین کارایی هستند و از لحاظ معیارها، Accuracy مدل‌های شبکه عصبی و درخت تصمیم دارای بیشترین دقت هستند. با این وجود این دو روش از لحاظ معیارهای FNrate و Recall، کارایی ضعیف‌تری نسبت به مدل رگرسیون لجستیک دارند. با توجه به اینکه هدف در پیش‌بینی ریزش مشتری، شناسایی مشتریانی است که ریزش آنها محتمل است، از اینرو ممکن است برای مدیران شرکت‌ها معیارهای FNrate و Recall، مهم‌تر از سایر معیارهای سنجش دقت باشند، که در اینصورت مدل رگرسیون لجستیک گزینه‌ی مناسبی برای پیشنهاد به مدیران شرکت‌ها خواهد بود.

مراجع

Blake, C. L. and Merz, C. J. (1998), Churn Data Set, UCI Repository of Machine Learning Databases, <http://www.ics.uci.edu/~mllearn/MLRepository.html>. University of California, Department of Information and Computer Science, Irvine, CA.

- Chawla, N.V. and Bowyer, K.W. and Hall, L.O. and Kegelmeyer, W. Ph. (2002), SMOTE: Synthetic Minority Over-sampling Technique, *Journal of Artificial Intelligence Research*.
- Garcia, V. and Sanchez, J. S. and Mollineda, R. A. and Alejo, R. and Sotoca, J. M. (2007), The class imbalance problem in pattern classification and learning, II *CongresoEspañol de Informatica*.
- Hastie, T. and Tibshirani, R. and Friedman, J. (2009), *The Elements of Statistical Learning*, Springer Series in Statistics ,<http://www.springer.com/series/692>.
- Keramati, A. and Ardabili S.M.S. (2011), Churn analysis for an Iranian mobile operator, *Telecommunications Policy*, **35**, 344-356.
- Mozer, M. C. and Dodier, R. and Colagrosso, M. D. and Guerra-Salcedo, C. and Wolniewicz, R. (2002), Prodding the ROC Curve: Constrained Optimization of Classifier Performance Advances, *Neural Information Processing Systems 14*, MIT Press.
- Rashid, T. (2009), Classification of Churn and non-Churn Customers for Telecommunication Companies Tarik , *International Journal of Biometrics and Bioinformatics (IJBB)*, Volume 3, Issue 5.
- Sharma, A. and Panigrahi, P.K. (2011), A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services, *International Journal of Computer Applications* (0975 - 8887), Volume 27- No.11.

Archive of SID