

داده کاوی دانشجویان انصرافی دانشگاه پیام نور استان تهران  
به منظور افزایش نرخ ماندگاری (جلوگیری از رویگردانی مشتری)

\*سید علی اکبر احمدی<sup>۱</sup>، تورج خیراتی کازرونی<sup>۲</sup>

۱. استاد گروه مدیریت دولتی، دانشگاه پیام نور، تهران، ایران.

۲. کارشناسی ارشد گروه مدیریت فناوری اطلاعات، دانشگاه تهران، تهران، ایران.

تاریخ دریافت: (۹۷/۸/۷) تاریخ پذیرش: (۹۷/۹/۱۲)

Data Mining Withdrawal of the Students of Payame Noor University in Tehran State to Increase Student Retention Rate (Preventing Customer Rejection)

\*Sayed Ali Akbar Ahmadi<sup>1</sup>, Toraj Khairati Kazerooni<sup>2</sup>

1. Professor of Department of Public Administration, Payame Noor University, Tehran, Iran.

2. Ph.D of Department of Information Technology Management, Tehran of University, Tehran, Iran.

Received: ((29/10/2018) Accepted: (3/12/2018)

Abstract

The challenges facing educational institutions, especially nonprofit higher education institutions, are to earn money to meet their goals. Student withdrawal is in the opposite direction. By identifying students as opt-outs, preventive and supportive policies can be anticipated to prevent a reduction in the image and hopes to attract revenues. This research is aimed at identifying students at risk by using the Data Mining Data of the Student attention Center of Payame Noor University of Tehran during the years 2013-16. Data were extracted from the education system. Of the 20 potentially effective attributes, 92% accuracy model was identified. In the model, six independent characteristics (age, group, grade, probation, and gender) and an associated attribute (term) were identified and subsequently the degree of importance of the attributes involved in the withdrawal and their relationship with each other was determined. Risk of withdrawal and risk ranking table for different terms were presented. Findings indicate that age is the most important factor. The bachelor degree is between the ages of 18-21 in the senior age group of 22-26 and the PhDs 31 to 29 in the most risky groups. In terms of academic and postgraduate degrees, they are the most risky period in the formal and continuing education programs of the research course. Attention rates for students aged 19 & 20 are about 50% in the third semester.

Keywords

Student Withdraws, Retention Students, Prevent Customer Churn, Baizian Network, SLR Algorithms.

چکیده

از چالش‌های پیش‌روی مؤسسات آموزشی بالأخص مؤسسات آموزش عالی غیرانتفاعی کسب درآمد جهت تحقق اهداف است اما انصراف دانشجو در نقطه مقابل قرار دارد. با شناسایی دانشجویان انصرافی می‌توان با اتخاذ سیاست‌های پیشگیرانه و حمایتی از کاهش وجهه مؤسسه جلوگیری و امیدوار به جذب درآمد مورد انتظار شد. این تحقیق با استفاده از داده‌های اطلاعات دانشجویان انصرافی دانشگاه پیام نور استان تهران طی سال‌های ۹۱ تا ۹۴ قصد دارد دانشجویان در معرض خطر را شناسایی کند. داده‌ها از سامانه آموزش استخراج و از ۲۰ صفت احتمالی مؤثر در انصراف مدلی با دقت ۹۲٪ شناسایی شد. در مدل مذکور ۶ مشخصه مستقل (سن، گروه، مقطع و دوره تحصیلی، مشروطی و جنسیت) و یک مشخصه وابسته (سنوات) شناسایی و متعاقباً درجه اهمیت مشخصه‌های دخیل در انصراف و ارتباط آن‌ها با یکدیگر تعیین شد. احتمال خطر انصراف (ریسک ریزش) اولویت‌بندی و جدول خطر احتمال برای ترم‌های مختلف ارائه شد. یافته‌ها حکایت از شناسایی سن به‌عنوان مهم‌ترین عامل دارد. از نظر سنی در کارشناسی دسته سنی ۱۸-۲۱ در ارشد ۲۶-۲۲ و در دکتری ۳۱-۲۹ پرخطرترین گروه‌ها هستند. از لحاظ دوره تحصیلی در کارشناسی و دکتری دوره رسمی و در ارشد دوره آموزشی پژوهشی رسمی پرخطرترین دوره می‌باشند. نرخ انصراف برای دانشجویان ۱۹ و ۲۰ سال در ترم سوم تقریباً ۵۰ درصد است.

واژه‌های کلیدی

انصراف دانشجو، حفظ دانشجو، رویگردانی مشتری، شبکه بی‌زین، پاسخ خود یادگیرنده.

## مقدمه

طی دهه‌های گذشته اشتیاق به داشتن تحصیلات عالی، کمبود ظرفیت آموزش عالی دولتی در پاسخگویی به این اشتیاق در کنار کاهش اعتبارات بخش دولتی موجب شکل‌گیری بخش غیرانتفاعی آموزش عالی شد. با ظهور مؤسسات غیرانتفاعی آموزش عالی همچون دانشگاه آزاد و کسب موفقیت‌ها و سود حاصل در این زمینه، بخش خصوصی تمایل بیشتری را جهت ورود به این عرصه نشان داد. مزایای این ورود نه تنها معطوف به کسب درآمد بود بلکه می‌توانست کشور را در تحقق اهداف برنامه‌های توسعه‌ای یاری داده و موجبات افزایش کمی و کیفی دانش‌آموختگان آموزش عالی را فراهم کند. یکی از این مؤسسات دانشگاه پیام نور است. یکی از چالش‌ها و تهدیدهای پیش‌روی مؤسسات آموزش عالی (چه دولتی و چه خصوصی) از جمله دانشگاه پیام نور انصراف دانشجویان است. اتخاذ استراتژی نگه داشت دانشجویان یک سرمایه‌گذاری ارزشمند جهت تداوم کسب درآمد و تضمین سود در آینده است. از این‌رو، مؤسسات آموزش عالی به این نتیجه رسیده‌اند جلوگیری از انصراف دانشجویان جهت موفقیت بلندمدت کسب‌وکار حیاتی است. «هم‌اکنون بحث بر سر کسب مزیت رقابتی در سایه بررسی داده‌های مخازن مختلف اطلاعاتی و کشف حقایق و دانش نهفته در دل آنان است. در آینده جمع‌آوری داده‌های مشتریان یک منبع مزیت رقابتی فزاینده خواهد بود (نشریه بیزینس دانشگاه هاروارد، ۲۰۱۵)». این تحقیق داده‌کاوی را به منظور کشف دانش نهفته در اطلاعات ذخیره شده دانشجویان در سامانه آموزش به مدد گرفته است. صرفه اقتصادی و آموزه‌های مدیریت حکم به حفظ دانشجویان دارد از سویی دیگر به منظور کاهش هزینه‌های حفظ دانشجویان (مشتری) می‌باید این سیاست‌ها معطوف به دانشجویانی باشد که با احتمال بیشتری قصد انصراف دارند. باید به نوعی قادر به پیش‌بینی رفتار آتی آنان بود تا با اتخاذ سیاست‌های مداخله‌جویانه مناسب از انصراف آنان ممانعت به عمل آورد. بنابراین، انصراف دانشجویان از مهم‌ترین چالش‌های پیش‌روی مؤسسات آموزش عالی چه دولتی چه غیردولتی است. شناسایی عوامل انصراف دانشجویان دانشگاه پیام نور استان تهران در سایه ذخیره و بازیابی اطلاعات دانشجویی هدف تحقیق است. هدف این تحقیق آن است که مشخص شود کدام یک از عواملی همچون سن، جنسیت، بومی بودن، مقطع تحصیلی، سنوات، دوره، گروه و رشته تحصیلی بر انصراف تأثیر گذاشته و کدام یک بیشترین

سهم را دارد. بدین منظور داده‌های مناسب فردی و تحصیلی دانشجویان جمع‌آوری و تجزیه و تحلیل شد تا ابتدا عوامل انصراف و سپس دسته‌های دانشجویان در معرض خطر زیاد انصراف شناسایی شده تا با اعمال سیاست‌های مداخله‌جویانه بازدارنده بتوان بقای آنان را تضمین کرد.

## پیشینه‌های پژوهش

تعداد دانشجویان از پارامترهای مؤثر در برنامه‌ریزی مراکز آموزشی است. پذیرش دانشجویان به‌منزله کسب درآمد و انصراف دانشجویان ترمزی برای آن است که اعتبار مرکز را نیز زیر سؤال می‌برد. انصراف، سرمایه‌گذاری آموزش عالی به امید برداشت محصول (دانش‌آموخته) را با مشکل روبرو می‌کند و حتی موجب افت محصول در سال‌های بعد نیز خواهد شد. دوام مؤسسات شهریه پرداز به حضور دانشجویان وابسته است و انصراف برای آنان از اهمیت ویژه‌ای برخوردار است. بنابراین، حفظ دانشجویان هم به معنای کسب درآمد بیشتر و هم به معنای حفظ وجهه موسسه است. در نگاه کلان، شناسایی الگوهای انصراف دانشجویان در آموزش عالی کشور و از نگاه خاص، شناسایی الگوی انصراف دانشگاه پیام نور می‌تواند منجر به شناسایی دسته‌های در معرض خطر انصراف شده تا با اتخاذ سیاست‌های پیشگیرانه و حمایتی زمینه ماندگاری دانشجویان را فراهم کرد.

دو رویکرد پیشگیرانه هدفمند و غیرهدفمند وجود دارد. پیشگیری هدفمند شامل دو نوع انفعالی (واکنشی) و پیشگیرانه (فعال) است (بارز و ون دن پل، ۲۰۰۷) در رویکرد انفعالی سازمان تا زمانی که علائم قطع ارتباط (انصراف) نمایان نشده منتظر می‌ماند. اما در رویکرد پیشگیرانه سازمان تمایل دارد پیش از قطع ارتباط عوامل یا علل قطع ارتباط شناسایی شده و با اتخاذ سیاست‌های مناسب تلاش دارد آنان را به ماندن ترغیب کند. رویکرد این مقاله پیشگیری هدفمند فعال است که تلاش دارد با تحلیل اطلاعات دانشجویان انصرافی ذخیره شده در بانک اطلاعات دانشجویی، ویژگی‌هایی که منجر به انصراف می‌شود را شناسایی کرده و با اتکا به آنها بتواند دانشجویان در معرض خطر انصراف را مشخص کرده و با اعمال اقدامات مداخله‌جویانه زمینه بقا آن‌ها را فراهم کند.

تحقیقات بسیاری در زمینه شناسایی عوامل مؤثر در انصراف، رتبه‌بندی و دسته‌بندی دانشجویان در معرض خطر انجام شده است. عمده‌ترین این تحقیقات به شرح زیر است: دانشگاه آلاباما پی برد که سفرهای متعدد (درون یا

شبکه‌های عصبی و تجزیه و تحلیل تشخیص خطی استفاده شد (ساپربای و همکاران<sup>۳</sup>، ۲۰۰۶).

در تحقیق یو و همکاران نرخ حفظ دانشجو با استفاده از داده‌های دانشجویان سال دوم و سوابق سال اول آنان ارزیابی گردید. داده‌های هویتی، جمعیت‌شناسی، سوابق تحصیلی قبل از دانشگاه، برنامه تحصیلی فعلی و پیشرفت تحصیلی لحاظ شده بودند. یافته‌ها نشان می‌دهد که حفظ دانشجو با ساعات مطالعه دروس دانشگاه و عوامل مکانی مانند محل اقامت گره‌خورده است. ارتباط جالب توجه مابین عوامل مذکور و مالی با عواملی همچون مسائل عاطفی محرز شد (یو<sup>۴</sup>، ۲۰۰۷).

پیتمن از شبکه‌های عصبی، رگرسیون لجستیک، طبقه‌بندی بیز و درخت تصمیم‌گیری جهت پیش‌بینی انصراف استفاده کرد (پیتمن<sup>۵</sup>، ۲۰۰۸).

دکر و همکاران با هدف پیش‌بینی انصراف دانشجو طی سال‌های ۲۰۱۰-۲۰۰۹ از الگوریتم‌های درخت تصمیم‌گیری، طبقه‌بندی بیز، مدل لجستیک، یادگیری مبتنی بر قواعد و جنگل تصادفی استفاده کردند (دکر<sup>۶</sup>، ۲۰۰۹: ۴۱).

ژانگ و همکاران به دنبال یافتن راهی جهت افزایش نرخ حفظ دانشجو بودند. برای این کار از الگوریتم‌های بیز، ماشین‌های بردار پشتیبان (SVM) و درخت تصمیم‌گیری استفاده شد. نتیجه حاصل نشان داد که تحلیل بیز نسبت به دو الگوریتم دیگر از دقت بالاتری در پیش‌بینی برخوردار بود (ژانگ و همکاران<sup>۷</sup>، ۲۰۱۰).

بار دیگر یو و همکاران تلاش کردند با استفاده از داده‌کاوی عوامل مؤثر بر حفظ دانشجو را شناسایی کنند. از الگوریتم‌های درخت تصمیم‌گیری، رگرسیون چند متغیره تطبیقی (MARS) و شبکه‌های عصبی استفاده شد. عواملی همچون اقامت و قومیت به‌عنوان عوامل حیاتی در حفظ دانشجو شناسایی شدند (یو و همکاران<sup>۸</sup>، ۲۰۱۰).

تحقیق زلاتکو کووچیچ متغیرهای اجتماعی جمعیتی مانند سن، جنس، نژاد، تحصیلات، وضعیت کار و ناتوانی‌های خاص و محیط را بررسی و مطالعه کرد. با استفاده از الگوریتم‌های طبقه‌بندی و درخت رگرسیون با بیش از ۶۵ درصد دقت نرخ انصراف را پیش‌بینی کرد (زلاتکو کووچیچ<sup>۹</sup>، ۲۰۱۰).

برون‌شهری) احتمال انصراف را افزایش می‌دهد. در نتیجه در راستای حفظ دانشجو زمینه اقامت تمام دانشجویان سال اول در محوطه دانشگاه فراهم شد و برای دانشجویان در معرض خطر برنامه‌ها و گردهمایی‌های خاص اجرا شد (بلا آموزش عالی آمریکا<sup>۱</sup>، ۲۰۱۲).

دانشگاه مهندسی سانی در بوفالو سیاست حمایتی مبنی شرکت در جلسات خصوصی تخصصی (کمک آموزشی پیش‌دانشگاهی) برای ورودی‌های جدید (بلاگ آموزش عالی آمریکا<sup>۲</sup>، ۲۰۱۲).

استراتژی حفظ دانشجویان جدید دانشگاه تیفین شامل تخصیص مربی خصوصی برای دانشجویان با نرخ ریسک بالا، ایجاد یک رئیس اداره حفظ دانشجو به‌منظور نظارت بر امور دانشجویان در معرض خطر با ریسک متوسط و اطلاع‌رسانی کلاس‌های فوق‌برنامه از طریق ارسال یک پیام پست الکترونیکی برای دانشجویان با نرخ ریسک پایین بود. دانشگاه از طریق اجرای این تلاش‌ها استراتژی حفظ دانشجو خود را بهبود بخشید. نرخ حفظ دانشجو طی پنج سال از ۵۱ درصد به ۶۳ درصد افزایش یافت (بلاگ آموزش عالی آمریکا<sup>۳</sup>، ۲۰۱۲).

تحقیق مینایی - بیگدلی تلاش کرد با استفاده از یک مجموعه از عامل‌های دسته‌بندی شده بر روی داده‌های بر خط دانشگاه ایالتی میشیگان، دانشجویان در معرض خطر انصراف را شناسایی کند آنان موفق شدند با ارائه مشاوره مطلوب توسط مربیان و مشاوران ماندگاری را به‌طور چشمگیری افزایش دهند (مینایی - بیگدلی<sup>۴</sup>، ۲۰۰۳: ۱-۶).

وانگ جهت حفظ دانشجو از رگرسیون لجستیک، شبکه عصبی و درخت تصمیم‌گیری استفاده کرد. نتیجه‌گیری بر کیفیت میزان یادگیری دانشجویان در دوره دبیرستان به استناد معدل نمره رشته و آزمون استعداد تحصیلی تأکید داشت. این موارد به‌عنوان مهم‌ترین عوامل حفظ دانشجو شناسایی شدند. انگیزه‌های شغلی و علایق دانشجویان از دیگر عوامل بودند. این تحقیق علاوه بر طرح پیشنهادهایی برای غلبه بر مشکل انصراف دانشجو یک برنامه حفظ و نگهداری نیز ارائه کرد (وانگ<sup>۵</sup>، ۲۰۰۵).

ساپربای و همکاران دانشجویان را با توجه به احتمال مواجهه با خطر ترک تحصیل به ۳ دسته کم‌خطر، متوسط‌الخطر و پرخطر تقسیم کردند. در این تحقیق از الگوریتم‌های طبقه‌بندی همچون درخت تصمیم‌گیری، جنگل تصادفی،

3. Superby, Vandamme & Meskens

4. Yu

5. Pittman

6. Dekker

7. Zhang, Oussena, Clark, & Kim

8. Yu, Digangi, Jannasch-pennell & Kaprolet

9. Zlatko Kovacic

1. Higher Education Blog

2. Vang

است مسئولان دانشگاهی وارد شده و رفع مشکل می‌کنند. سعی می‌کنیم با مشاوره علت‌جویی کرده و تصمیم آنها را نسبت به انصراف تغییر دهیم (خبرگزاری تابناک، ۱۳ تیر ۸۹).  
 دکتر عبدالله شفیع‌آبادی معاون آموزشی دانشگاه علامه طباطبایی نیز در گفتگو با مهر گفت: طرح بومی‌گزینی شیوه مناسبی برای کاهش تعداد اخراجی‌ها و انصرافی‌های دانشگاهها است و اجرای این طرح می‌تواند از تعداد انصرافی یا اخراجی‌ها بکاهد. ایشان با بیان اینکه «باید انتخاب رشته درست و اصولی صورت گیرد» عامل اصلی انصراف و اخراج دانشجویان را انتخاب رشته به‌شیوه نادرست عنوان کرد و گفت: مصیبتی که بسیاری از دانشجویان در بدو ورود گرفتار آن می‌شوند تنها به دلیل تن دادن به چشم و هم چشمی‌ها و مشاوره‌های نادرست قبل از انتخاب رشته است. دانشجویی که بدون شناخت توانمندی و علایق خود، روزها را صرف مطالعه برای شرکت در کنکور می‌کند پس از قبولی در دانشگاه متوجه می‌شود که هیچ علاقه‌ای به رشته قبولی خود ندارد (خبرگزاری تابناک، ۱۳ تیر ۸۹).

رئیس دانشگاه بوعلی همدان نیز در گفتگو با مهر دلایل انصراف دانشجویان در دوره‌های روزانه را عدم علاقه به رشته قبولی و دوری از شهرستان محل سکونت و در دوره‌های شبانه تأمین هزینه بیان کرد. علی محمد زلفی گل ادامه داد: در دانشگاه بوعلی همدان بیشتر دانشجویان دوره شبانه انصراف می‌دهند (خبرگزاری تابناک، ۱۳ تیر ۸۹).

دکتر زلفی گل عامل مهم در جلوگیری از انصراف و اخراج را ارتباط خانواده دانشجویان با دانشگاه دانست و افزود: ارتباط خانواده دانشجویان و دانشگاه می‌تواند مشکلات را به حداقل برساند. خانواده‌ها باید مثل دبیرستان پیگیر روند تحصیلی فرزندان خود باشند. همچنین باید در انتخاب رشته مشاوره دقیق کرد تا رشته مطلوب انتخاب شود. راهنمایی‌های مشاور نیز باید با توجه به سوابق قبلی فرد صورت گیرد. بیشتر شبانه‌ها انصراف می‌دهند. دلایل انصراف دانشجویان در دوره‌های روزانه را عدم علاقه به رشته قبولی و دوری از شهرستان محل سکونت و در دوره‌های شبانه را مشکلات ناشی از تأمین هزینه دانست (خبرگزاری تابناک، ۱۳ تیر ۸۹).

دکتر مرتضی زرگر شوشتری رئیس دانشگاه شهید چمران اهواز نیز به مهر گفت: رشته‌های کم‌طرفدار اخراجی یا انصرافی بیشتری دارند (خبرگزاری تابناک، ۱۳ تیر ۸۹).

دکتر محمد اسماعیل همدانی گلشن رئیس دانشگاه صنعتی اصفهان معتقد است دانشجویانی که اخراج یا انصراف می‌دهند کیفیت آموزشی مطابق با مقررات آموزشی را دارا نیستند. این

دلن با استفاده از چند تکنیک داده‌کاوی مدل ترک تحصیل دانشجویان سال اول را ارائه کرد. الگوریتم‌های به کار برده شده شامل درخت تصمیم‌گیری، SVM، شبکه‌های عصبی و رگرسیون لجستیک بودند. از میان چهار مدل پیش‌بینی مورد استفاده در این تحقیق SVM بهترین نتیجه را نشان داد پس از آن به ترتیب درخت تصمیم، شبکه‌های عصبی و رگرسیون لجستیک قرار گرفتند (دلن، ۲۰۱۰: ۴۹۸).

تحقیق بوگارد و همکاران بر روی دانشجویان دانشکده کنتاکی غربی (آمریکا) بود. از مدل تجمیعی مرکب از رگرسیون لجستیک، درخت تصمیم‌گیری و شبکه عصبی استفاده شد (بوگارد، ۲۰۱۲).

تحقیقات یو و همکاران (۲۰۰۷)، پیتمن (۲۰۰۸) و بوگارد و همکاران (۲۰۱۳) تأکید و تمرکز بیشتری بر استفاده از داده‌های آموزشی ذخیره شده در پایگاه‌های داده در مقایسه با داده‌های جدید جمع‌آوری شده (از طریق پرسشنامه) دارند. از نظرسنجی به‌مراتب کمتر جهت جمع‌آوری داده‌ها استفاده شده است. درواقع، از نظرسنجی -در سطح محدودی- به‌عنوان روش جمع‌آوری داده استفاده شده است. پژوهش وانگ در سال ۲۰۰۵ آخرین تحقیق از این نوع بود که در آن به‌ندرت از اطلاعات موجود استفاده شده بود. به‌عنوان نتیجه‌گیری کلی می‌توان گفت داده‌های آموزشی مهم‌ترین منبع داده برای تحقیقات داده‌کاوی آموزشی بالأخص داده‌کاوی ترک تحصیل است.

اغلب تحقیقات انجام شده برای روی دانشجویان انصرافی درخصوص یافتن علل انصراف است که براساس اطلاعات حاصل از پرسشنامه به دست آمده‌اند. در زیر ضمن اشاره به برخی از آنها، نظر مسئولان ده دانشگاه برتر ایران در نشست آنان به شرح زیر است:

دکتر سید مهدی قمصری معاون آموزشی دانشگاه تهران بیان نمودند: در دانشگاه تهران انصراف از تحصیل در صورتی رخ می‌دهد که دانشجو شرایط بهتری برای ادامه تحصیل درجایی دیگر داشته باشد (خبرگزاری تابناک، ۱۳ تیر ۸۹).

دکتر علی مقداری معاون آموزشی دانشگاه صنعتی شریف اظهار کردند افرادی که انصراف می‌دهند بیشتر دچار مشکلات خانوادگی یا اقتصادی هستند و یا فرصت بهتری برای تحصیل در خارج از کشور را می‌یابند. دکتر مقداری درخصوص برخورد دانشگاه با انصراف دهندگان نیز گفت: در زمان‌هایی که لازم

عضو هیئت‌امنا دانشگاه پیام نور در ادامه بودجه این دانشگاه را ۱۰ هزار میلیارد ریال ذکر و بیان کرد: حدود سه هزار و ۵۰۰ میلیارد ریال از این مبلغ صرف هزینه‌های پرسنلی می‌شود. وی افزود: هم‌اکنون خدماتی که دانشجویان از سوی کارکنان دانشگاه پیام نور در بخش‌های رفاهی و آموزشی دریافت می‌کنند کمتر از هزینه‌ای است که برای پرسنل پرداخت می‌شود (پایگاه خبری دانشجویان پیام نور، اردیبهشت ۱۳۹۴).

### روش‌شناسی پژوهش

این تحقیق کاربردی و از نوع توصیفی - تحلیلی است که از داده‌های کمی - کیفی بر مبنای متد کریسپ داده‌کاوی<sup>۱</sup> بر روی اطلاعات فردی و تحصیلی دانشجویان دانشگاه پیام نور استان تهران انجام گرفته است. مجموعه داده‌ها شامل اطلاعات دانشجویان انصرافی سال‌های تحصیلی ۱۳۹۰-۹۱ لغایت ۵۹-۱۳۹۴ است. به‌منظور تحلیل و مدل‌سازی از نرم‌افزار کلمنتاین<sup>۲</sup> از مجموعه SPSS<sup>۳</sup> استفاده شد. در کانون اهداف تحقیق، کنترل و کاهش نرخ ریسک ریزش، افزایش و یا حفظ اعتبار مؤسسه آموزش عالی، افزایش درآمد، جلوگیری از اتلاف منابع بالأخص منابع انسانی قرار دارد. سؤال‌های طرح شده در این تحقیق عبارت‌اند از:

- آیا می‌توان عوامل مؤثر در انصراف را شناسایی کرد؟
- آیا می‌توان از ترم (های) خاصی نام برد که نقطه بحرانی در انصراف باشد؟
- آیا نتایج حاصل می‌تواند به مدیران ارشد در برنامه‌ریزی‌های آموزشی کمک کند؟
- فرضیه‌های زیر در نظر گرفته شده‌اند:
- دانشجویان در سنین مختلف در قبال انصراف رفتاری متفاوت نشان می‌دهند.
- رفتار دانشجویان انصرافی مقاطع تحصیلی با یکدیگر متفاوت است.
- انصراف در گروه‌های مختلف تحصیلی متفاوت است.
- سال اول مرحله بحرانی برای خروج از تحصیل است.
- جنسیت از عوامل عمده تأثیرگذار بر انصراف است.
- بومی/ غیربومی بودن دانشجو یکی از عوامل عمده تأثیرگذار بر انصراف از تحصیل است.

دانشجویان عمدتاً امکان وفق خود با محیط دانشگاه را نداشته‌اند و یا انگیزه لازم برای ادامه تحصیل در رشته‌ای را که قبول شده‌اند ندارند. وی از سوی دیگر به مشکلات شخصی و یا فرهنگی به‌عنوان دلایل دیگر انصراف یا اخراج از تحصیل اشاره کرد و گفت: نمی‌توان با قطعیت گفت تعداد دانشجویان اخراجی در کدام رشته‌ها بیشتر است زیرا عدم انگیزه برای تلاش در هر رشته‌ای ممکن است بروز پیدا کند. اما آنچه مسلم است در رشته‌هایی که طرفدار کمتری دارد عدم انگیزه بیشتر مشهود است (خبرگزاری تابناک، ۱۳ تیر ۸۹).

در تحقیقی که از فاطمه گلشنی و فریبا حسنی با عنوان «محاسبه و تحلیل نرخ افت تحصیلی و عوامل مؤثر بر آن» بر روی جامعه آماری مشتمل بر ۴۵۰۸ دانشجوی شاغل به تحصیل سال تحصیلی ۷۹-۱۳۷۸ از شهرهای تهران، تبریز، شیراز، اهواز، اصفهان، مشهد، کرمان و گیلان انجام شد. جنسیت از مهم‌ترین عوامل در انصراف تشخیص داده شد و کلیه تحلیل‌ها با توجه به آن انجام شد. در مقطع کارشناسی، بیشترین نرخ انصراف در نیمسال اول رخ داده بود. در مقطع کارشناسی ارشد بیشترین نرخ انصراف در نیمسال دوم رخ داده بود (انتشارات مؤسسه پژوهش و برنامه‌ریزی آموزش عالی، پاییز ۱۳۸۰).

یافته‌های مریم مصباحی و مریم صابری نشان داد که کاهش علاقه به رشته تحصیلی، مشکلات شخصی، ناامیدی به آینده شغلی و بازار کار، حجم منابع درسی و سختی آزمون‌ها، شهریه، نگرش منفی به دانشگاه پیام نور و مدرک آن، امکانات و خدمات آموزشی دانشگاه و قبولی در سایر دانشگاه‌ها بیشترین تأثیر را در انصراف دانشجویان از دانشگاه دارند (همایش بین‌المللی افق‌های نوین در علوم مدیریت و حسابداری، اقتصاد و کارآفرینی، ۱۳۹۵).

### انصراف در دانشگاه پیام نور

در اردیبهشت ماه ۱۳۹۴ دکتر سید احمد میرشکرایی استاد تمام و عضو هیئت‌امنا دانشگاه پیام نور در پایگاه خبری دانشجویان پیام نور (پیام‌نورنا) اعلام داشتند شمار دانشجویان این مرکز آموزش عالی طی پنج سال گذشته ۲۸ درصد کاهش یافته است. در ادامه بیان کردند که سال ۱۳۸۹ تعداد دانشجویان دانشگاه پیام نور یک‌میلیون و ۱۰۰ هزار نفر بود که امسال به ۸۰۰ هزار دانشجو کاهش یافته است. وی گفت: آمار انصراف از تحصیل در دانشگاه پیام نور رو به افزایش بوده و ارائه نشدن خدمات کافی در بخش‌های آموزش، پژوهش و رفاهی دلسردی دانشجویان را رقم زده است.

1. Crisp\_DM (Croos-Industry Standard Process for Data Mining)  
2. Celemantine  
3. Statistical Package for Social Science

آنها قادر به ارائه دسته‌بندی دانشجویان انصرافی کرده و برای هر دسته احتمال انصراف را پیش‌بینی کند.

### یافته‌های پژوهش

#### مرحله اول: یافتن بهترین الگوریتم جهت رتبه‌بندی پیش‌بینی‌کننده‌های انصراف

شبکه‌های بیزین کاربر را قادر می‌سازد که یک مدل احتمالی با ترکیب مشاهدات بسازد (جمال شهری، ۱۳۹۲) شکل ۱ نحوه پیاده‌سازی الگوریتم شبکه خانواده بیزین را نشان می‌دهد. داده‌ها به دو مجموعه داده‌های یادگیری (۸۰٪ کل دانشجویان انصرافی) و مجموعه داده‌های تست (۲۰٪ کل دانشجویان انصرافی) تقسیم شدند. الگوریتم سعی می‌کند با تحلیل داده‌های مجموعه یادگیری نحوه رفتار دانشجویان را فرا گرفته و برای اینکه میزان موفقیت خود را بسنجد برای هر دانشجو با توجه به متغیرهای مستقل او، یک پیش‌بینی برای تعداد سنوات انجام می‌دهد و در انتها با مقایسه این پیش‌بینی با مقادیر واقعی دقت مدل خود را می‌سنجد. از شبکه‌های بیزین، الگوریتم TAN که بالاترین دقت را داشت انتخاب شد. صفات سن، گروه، تعداد مشروطی، دوره تحصیلی، مقطع و جنسیت به‌عنوان پیش‌بینی‌کننده‌های انتخاب شدند. جهت بهبود مدل TAN از الگوریتم‌های بهبود یافته مارکوف و مارکوف FS استفاده شد دقت در مجموعه یادگیری افزایش یافت اما دقت داده‌های تست کم بود. برای افزایش دقت مجموعه داده‌های تست، اقدام به تولید الگوریتم‌های ترکیبی شد. همان‌گونه که در جدول ۱ دیده می‌شود خروجی هر الگوریتم ورودی الگوریتم دیگر می‌شود. این کار با توجه به اطلاعات جدول ۱ باعث بهبود مدل شد.

جدول ۱. مقایسه دقت الگوریتم‌های شبکه بیزین

الگوریتم	مجموعه داده یادگیری		مجموعه داده تست	
	درصد پیش	درصد پیش	درصد پیش	درصد پیش
TAN	۶۱/۵۹	۳۸/۰۵	۳۶/۴۳	۶۳/۵۷
مارکوف	۸۷/۲۹	۱۲/۷۱	۱۶/۷۹	۸۳/۲۱
مارکوف FS	۸۷/۲۹	۱۲/۷۱	۱۶/۷۹	۸۳/۲۱
ترکیبی TAN مارکوف	۹۱/۷۳	۸/۲۷	۵۰/۷۵	۴۹/۲۵

زیرا دقت مدل ترکیبی TAN به مارکوف برای مجموعه داده یادگیری، ۹۲ درصد است که بیشتر از دقت الگوریتم مارکوف (۸۷ درصد) است و همچنین دقت پیش‌بینی این الگوریتم در مجموعه داده تست ۵۱ درصد است که باز هم به مراتب بیشتر از دقت الگوریتم مارکوف (۱۷ درصد) است. بنابراین، دقت

اهداف تحقیق را می‌توان چنین عنوان کرد:

- دسته‌بندی دانشجویان در معرض خطر انصراف و شناسایی پرخطرترین‌ها.

- ارائه جدول احتمال انصراف برای دانشجویان با صفات مشترک.

- جلب نظر مدیران ارشد دانشگاه در خصوص اتخاذ سیاست‌های مداخله‌جویانه مناسب.

- زمینه‌سازی جهت ایجاد سیستم پشتیبان تصمیم‌گیری به‌منظور شناسایی و اقدام مناسب در قبال دانشجویان در معرض خطر انصراف.

قلمرو موضوعی تحقیق داده‌کاوی دانشجویان انصرافی دانشگاه پیام نور استان تهران با تمرکز بر حفظ دانشجویان است. قلمرو مکانی در این تحقیق دانشگاه پیام نور استان تهران است. قلمرو زمانی تحقیق یک بازه ۵ ساله مابین سال‌های تحصیلی ۹۱-۱۳۹۰ لغایت ۹۵-۱۳۹۴ است.

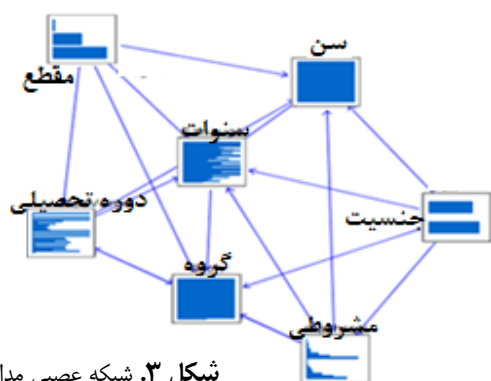
جامعه آماری تحقیق شامل اطلاعات فردی و تحصیلی دانشجویان ۹۰ تا ۱۳۹۴ بوده که از سیستم دانشجویی (گلستان) استخراج شده است. جامعه هدف شامل دانشجویان انصرافی جامعه آماری است. از مجموعه اطلاعات اخذ شده تعداد سنوات دانشجوی انصرافی به‌عنوان صفت هدف (متغیر وابسته) و متغیرهای مقطع تحصیلی، دوره تحصیلی، دانشکده، مرکز، گروه، رشته تحصیلی، جنسیت، بومی/غیربومی، وضعیت تأهل و سن به‌عنوان پیش‌بینی‌کننده (متغیرهای مستقل) در نظر گرفته شدند. برای بررسی فرضیه‌های مطرح شده تجزیه و تحلیل در چند مرحله به شرح زیر انجام خواهد شد:

مرحله اول: یافتن بهترین الگوریتم جهت رتبه‌بندی پیش‌بینی‌کننده‌های دخیل در انصراف به لحاظ اهمیت و ارائه جدول احتمال انصراف. برای نیل به این منظور ابتدا الگوریتم‌های مختلف شبکه‌های بیزین (بیزساده<sup>۲</sup> و الگوریتم‌های بیزین بهبود یافته شامل الگوریتم<sup>۳</sup> TAN، مارکوف<sup>۴</sup> و مارکوف<sup>۵</sup> FS) به لحاظ دقت در تحلیل و پیش‌بینی با یکدیگر مقایسه خواهند شد.

مرحله دوم: معرفی دسته‌های مختلف دانشجویان به لحاظ میزان مواجهه با خطر انصراف. در این مرحله جهت یافتن تشابه یا تمایز در رفتارهای دانشجویان در قبال انصراف و دسته‌بندی آنان از مدل پاسخ خود یادگیرنده استفاده شد. این الگوریتم سعی می‌کند با بررسی اطلاعات دانشجویان عواملی را شناسایی کند که توسط

1. Bayzien Networks
2. Naïve Bâyes
3. Tree Augmented Naïve Bâyes
4. Markov Blanket Estimation
5. Markov Feature Selection Estimation

- گروه بر سن و نوع ورود (دوره) تأثیر دارد. با دوره تأثیر متقابل دارد.
- سن تنها بر سنوات تأثیر مستقیم دارد.
- جنسیت بر سن، سنوات، تعداد مشروطی و گروه تأثیر دارد.
- تعداد مشروطی بر سن، سنوات و گروه تأثیر دارد.
- نوع ورود بر سن، سنوات و گروه اثر دارد.



شکل ۳. شبکه عصبی مدل

**ارتباط سن و مقطع در انصراف دانشجویان**  
 با استفاده از جدول احتمال مدل و پس از دسته‌بندی سن دانشجویان قوانین استخراج به شرح زیر است:  
 - ۸۳ درصد دانشجویان انصرافی کارشناسی بین ۱۸ تا ۲۶ و ۸۳ درصد دانشجویان انصرافی ارشد بین ۲۲ تا ۳۶ سال سن دارند.

- پرخطرترین سنین جهت انصراف دانشجوی کارشناسی ۱۸ تا ۲۱ است که دانشجویان ۲۰ ساله با نرخ ۲۱ درصد در رأس این گروه سنی قرار دارند.

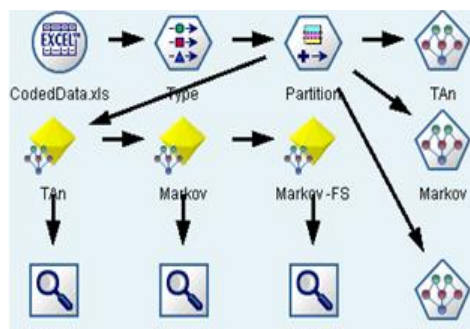
- پرخطرترین سنین جهت انصراف دانشجوی کارشناسی ارشد ۲۲ تا ۲۶ است که دانشجویان ۲۵ ساله با نرخ ۱۱ درصد در رأس این گروه سنی قرار دارند.

- پرخطرترین سنین جهت انصراف دانشجوی دکتری ۲۹ تا ۳۲ است که دانشجویان ۲۹ ساله با نرخ ۵۰ درصد در رأس این گروه سنی قرار دارند.

- ۵۴ درصد انصرافی‌ها مرد و ۴۶ درصد زن می‌باشند. در کل انصرافی‌ها تفاوت اندکی در احتمال انصراف زن و مرد وجود دارد. ۶۵/۳ درصد انصرافی‌ها کارشناسی، ۳۴/۲ درصد کارشناسی ارشد و ۰/۵ درصد دکتری تخصصی می‌باشند.

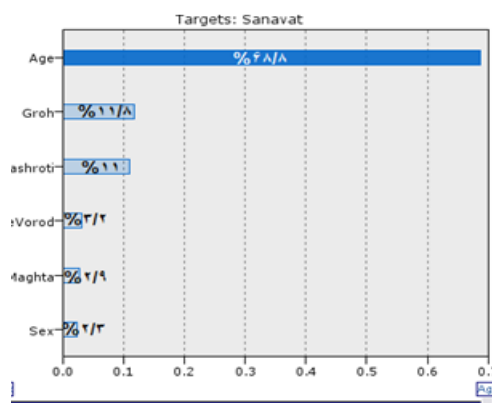
بیش از نیمی از دانشجویان کارشناسی با داشتن ۱ یا ۲ ترم انصراف می‌دهند. (دانشجوی ترم دوم و سوم) فقط ۱/۳۷ درصد

الگوریتم ترکیبی تن-مارکوف هم در داده‌های یادگیری و هم در داده‌های تست از الگوریتم مارکوف بیشتر است پس این مدل موجب بهبود شده و مدل دقیق‌تری را ارائه کرده است. در ادامه با توجه به انتخاب الگوریتم تن - مارکوف به‌عنوان دقیق‌ترین الگوریتم بررسی شد از لحاظ این الگوریتم درصد اهمیت هر یک از صفات پیش‌بینی‌کننده چه میزان است.



شکل ۱. اجرای الگوریتم شبکه بیزین

شکل ۲ درجه اهمیت متغیرهایی که مدل تن-مارکوف به‌عنوان صفات مهم در انصراف شناسایی کرده را نشان می‌دهد. سن با ۶۸/۸ درصد مهم‌ترین صفت در انصراف است. گروه تحصیلی، سنوات تحصیلی، دوره تحصیلی، مقطع تحصیلی و جنسیت در رتبه‌های بعدی قرار دارند.



شکل ۲. اهمیت متغیرها

شکل ۳ شبکه عصبی مدل و ارتباط متغیرها با یکدیگر را به تصویر می‌کشد. این مدل ارتباط پیش‌بینی‌کننده با صفت هدف را نیز نشان می‌دهد. پیکان‌ها به‌منزله سلسله اعصاب مدل می‌باشند. قوانین استخراج با توجه به این شکل به شرح زیر است:

- مقطع تحصیلی بر سن، سنوات، دوره تحصیلی و گروه تأثیر دارد.

انصرافی از دسته سنی خود و رتبه سوم در گروه علوم انسانی و در دسته ۱۸ تا ۲۲ با نرخ ۵۷ درصد انصرافی از دسته سنی خود است.

- رتبه اول انصراف مردان در گروه علوم فنی مهندسی و در دسته ۲۳ تا ۲۵ با نرخ ۶۳ درصد انصرافی از دسته سنی خود، رتبه دوم در گروه علوم فنی مهندسی و در دسته ۱۸ تا ۲۲ با نرخ ۶۰ درصد انصرافی از دسته سنی خود و رتبه سوم در گروه علوم فنی مهندسی و در دسته بیشتر از ۲۵ سال با نرخ ۵۶ درصد انصرافی از دسته سنی خود است.

### مرحله دوم - یافت مدل پیش‌بینی احتمال انصراف - مدل پاسخ خود یادگیرنده

این مدل برای هر دانشجو حداکثر ۵ حالت خاص (در صورت امکان) که منجر به انصراف خواهد شد را بررسی و بهترین مدل را انتخاب و ارائه خواهد کرد. مدل برای یادگیری از مجموعه داده یادگیری استفاده می‌کند. مجموعه داده یادگیری ۸۰ درصد مجموعه کل داده است. به عبارتی دیگر مدل ۱۱۰۹ دانشجو را بررسی کرده و به دنبال شناسایی صفتهایی (پیش-بینی‌کننده‌هایی) است که به استناد آنها بتوان نرخ احتمال انصراف را تعیین کرد.

id	Segment	Score	Cover (t)	Frequency	Probability
	All segments including Remainder		1,109	33	2.98%
1	mashrofi, NoeVorod mashrofi = 0 and NoeVorod = "A-R"	0.0	76	17	22.37%
2	Age Age = 25	0.0	66	5	7.50%
	Remainder		967	11	1.14%

شکل ۴. سنوات

شکل ۴ نمونه‌ای از مدل پاسخ خود یادگیرنده است. در اینجا مدل تعداد سنوات را صفر فرض کرده است مدل تلاش دارد روندی را کند و بر مبنای آن مدل و الگوی مناسب که توجیه‌کننده این رفتار باشد را تشخیص و برای اساس آن مقدار احتمال انصراف دانشجو را ارائه دهد. قوانین استنتاج از مدل جهت ترم‌های مختلف به شرح زیر است:

- ۲۲ درصد دانشجویان ارشد انصرافی دوره آموزشی رسمی با تعداد مشروطی صفر در ترم اول انصراف می‌دهند.

دانشجویان انصرافی کارشناسی در ترم اول (با صفر ترم) انصراف می‌دهند. آنان مترصد آن هستند که خود را با برنامه و محیط جدید آموزشی تطبیق دهند. در ترم دوم و سوم ۵۷ درصد آنان تشخیص می‌دهند نمی‌توانند خود را با محیط تطبیق دهند و از محیط جدید خارج می‌شوند (انصراف می‌دهند) از ترم چهارم به بعد تطابق‌پذیری بیشتر شده و روند انصراف کاهش می‌یابد. بنابراین، پرخطرترین دانشجویان کارشناسی در معرض انصراف دانشجویان ترم دوم و سوم هستند. برای دانشجویان انصرافی ارشد وضع به همین منوال است با کمی تفاوت. در کارشناسی انصرافی ترم سوم از دوم بیشتر است اما در ارشد برعکس است. اما به هر حال بازهم پرخطرترین دانشجویان ارشد در معرض انصراف دانشجویان ترم دوم و سوم هستند. درخصوص دانشجویان انصرافی دکتری باید اذعان داشت همه انصرافی‌ها در ترم دوم انصراف می‌دهند.

- ۸۱ درصد دانشجویان انصرافی کارشناسی بین ۲ تا ۴ ترم و ۸۳ درصد دانشجویان انصرافی ارشد بین ۲ تا ۴ ترم دارند.

- پرخطرترین ترم جهت انصراف دانشجوی کارشناسی ۱ تا ۲ ترم است که دانشجویان با ۲ ترم و با نرخ ۳۱ درصد در رأس این گروه سنی قرار دارند.

- پرخطرترین ترم جهت انصراف دانشجوی کارشناسی ۱ تا ۲ ترم است که دانشجویان با ۱ ترم و با نرخ ۳۴ درصد در رأس این گروه سنی قرار دارند.

- پرخطرترین ترم جهت انصراف دانشجوی دکتری ترم ۱ است.

### ارتباط انصراف با جنسیت، گروه تحصیلی و سن

با توجه به تعداد کم انصرافی (۱۳۸۹ دانشجو) و تعداد گروه و سن از رویکرد دسته‌بندی استفاده شد گروه‌های تحصیلی در قالب گروه‌های تحصیلی یونسکو دسته‌بندی شدند. بیشترین انصرافی زنان در گروه علوم انسانی با نرخ ۵۸ درصد و برای مردان در گروه فنی مهندسی با نرخ ۵۹ درصد است. شایان ذکر است برای زنان انصرافی در گروه هنر و علوم فنی مهندسی دیده نمی‌شود.

- رتبه اول انصراف زنان در گروه علوم انسانی و در دسته ۲۳ تا ۲۵ با نرخ ۶۲ درصد انصرافی دسته از سنی خود، رتبه دوم در گروه علوم انسانی و در دسته بیشتر از ۲۵ با نرخ ۶۰ درصد



- ۱۹ درصد دانشجویان انصرافی با بیش از ۳ ترم مشروطی، در ۱۹ ترم هشتم انصراف می‌دهند.
- ۱۲ درصد دانشجویان انصرافی با بیش از ۳ ترم مشروطی، در ترم نهم انصراف می‌دهند.
- ۹ درصد دانشجویان انصرافی با ۳ ترم مشروطی، در ترم نهم انصراف می‌دهند.
- ۹ درصد دانشجویان انصرافی با بیش از ۳ ترم مشروطی، در ترم دهم انصراف می‌دهند.

### بحث و نتیجه‌گیری

درخصوص سؤال اول تحقیق (شناسایی عوامل مهم در انصراف) نتایج حکایت از شناسایی یک مدل با دقت ۹۲ درصد با استفاده از الگوریتم شبکه‌های بیزین دارد. درجه و میزان اهمیت عوامل شناسایی شده از طریق مدل در جدول زیر نشان داده شده است.

جدول ۲. رتبه‌بندی پیش‌بینی‌کننده‌ها

شرح	رتبه				
	۱	۲	۳	۴	۵
پیش‌بینی‌کننده	سن	گروه	تعداد مشروطی	دوره تحصیلی	مقطع جنسیت
اهمیت (درصد)	۷/۸	۱/۸	۱۱	۳/۲	۲/۹
	۲/۳				

درخصوص سؤال دوم تحقیق (شناسایی ترم بحرانی در انصراف) می‌توان بیان داشت که ۷۰ درصد دانشجویان انصرافی کارشناسی و ارشد با ۰ تا ۳ ترم جزو پرخطرترین‌های انصراف هستند. کلیه انصرافی‌های دکتری ۱ ترم دارند. به لحاظ ترمی احتمال انصراف دانشجویان زن ۳ برابر مرد است. این نتایج تأییدی بر فرضیه چهارم تحقیق (سال اول تحصیلی مرحله بحرانی است) نیز است. به طوری که دانشجویان زن با ۱ ترم و مرد با ۲ ترم پرخطرترین دانشجویان در معرض خطر انصراف هستند.

فرضیه اول تحقیق درخصوص سن تأیید می‌شود زیرا که پرخطرترین سنین برای انصراف دانشجویان کارشناسی ۱۸ تا ۲۱ سال است (۶۱ درصد) که در رأس آن‌ها دانشجویان زن با ۱۹ و مرد با ۲۰ سال قرار دارند. پرخطرترین سنین برای انصراف دانشجویان کارشناسی ارشد ۲۲ تا ۲۶ سال است (۳۸ درصد) که در رأس آن‌ها دانشجویان ۲۶ سال قرار دارند. پرخطرترین سنین برای انصراف دانشجویان دکتری ۲۹ تا ۳۲ سال است (۱۰۰

- ۸ درصد دانشجویان انصرافی ۲۵ ساله، در ترم اول انصراف می‌دهند.
- ۵۶ درصد دانشجویان ارشد انصرافی دوره آموزشی پژوهشی رسمی، در ترم دوم انصراف می‌دهند.
- ۵۰ درصد دانشجویان انصرافی زن بدون مشروطی در ترم دوم انصراف می‌دهند.
- ۴۴ درصد دانشجویان ارشد انصرافی دوره آموزشی رسمی با ۰ یا ۱ مشروطی، در ترم دوم انصراف می‌دهند.
- ۳۷ دانشجوی کارشناسی انصرافی دوره رسمی بدون مشروطی، در ترم دوم انصراف می‌دهند.
- ۲۶ درصد دانشجویان انصرافی بدون مشروطی، در ترم دوم انصراف می‌دهند.
- ۵۲ درصد دانشجویان ارشد انصرافی دوره آموزشی پژوهشی فراگیر، در ترم سوم انصراف می‌دهند.
- ۵۰ درصد دانشجویان انصرافی با ۲۰ سال سن، در ترم سوم انصراف می‌دهند.
- ۴۷ درصد دانشجویان انصرافی با ۱۹ سال سن، در ترم سوم انصراف می‌دهند.
- ۴۱ درصد دانشجویان ارشد انصرافی دوره آموزشی فراگیر، در ترم سوم انصراف می‌دهند.
- ۲۲ درصد دانشجویان کارشناسی انصرافی دوره فراگیر، در ترم چهارم انصراف می‌دهند.
- ۳۱ درصد دانشجویان انصرافی با تعداد مشروطی ۲ یا ۳، در ترم چهارم انصراف می‌دهند.
- ۲۰ درصد دانشجویان انصرافی با ۲۰ سال سن، در ترم پنجم انصراف می‌دهند.
- ۱۶ درصد دانشجویان کارشناسی انصرافی دوره فراگیر، در ترم پنجم انصراف می‌دهند.
- ۲۵ درصد دانشجویان انصرافی با بیش از ۳ ترم مشروطی، در ترم ششم انصراف می‌دهند.
- ۲۴ درصد دانشجویان انصرافی با ۳ ترم مشروطی، در ترم ششم انصراف می‌دهند.
- ۱۴ درصد دانشجویان انصرافی با ۲۱ سال سن، در ترم ششم انصراف می‌دهند.

### پیشنهادها

باید ترم اول و دوم برای دانشجویان ارشد دوره آموزشی رسمی و ترم دوم کارشناسی مدنظر باشد و با توجه به اینکه تقریباً نیمی از دانشجویان ۱۹ و ۲۰ سال در ترم سوم انصراف می‌دهند و یا ۳۴ درصد دانشجویان ۲۱ سال در ترم ۵ یا ۶ انصراف می‌دهند، دانشگاه باید سیاست‌های بازدارنده خود شامل بسته‌های تشویقی، مشاوره‌ای و حمایتی را معطوف این دانشجویان نماید. دانشگاه می‌تواند به‌منظور استراتژی حفظ دانشجویان برنامه‌های حمایتی شامل تخصیص مربی خصوصی برای دانشجویان با نرخ ریسک بالا، ایجاد یک رئیس اداره حفظ دانشجویان به‌منظور نظارت بر امور دانشجویان در معرض خطر با ریسک متوسط و اطلاع‌رسانی کلاس‌های فوق‌برنامه از طریق ارسال یک پیام پست الکترونیکی برای دانشجویان با نرخ ریسک پایین باشد.

با توجه به یافته‌هایی که درخصوص ارتباط جنسیت، مقطع، دوره تحصیلی، تعداد مشروطی و تعداد سنوات و تعیین سه رتبه نخست در معرض خطر انصراف به دست آمد دانشگاه باید در مقطع کارشناسی بر روی دانشجوی رسمی تمرکز کند. دانشجوی زن با ۱ ترم و مرد با ۲ ترم پرخطرترین آنها هستند. هرچند در مقطع ارشد باید تمرکز روی دانشجوی آموزشی رسمی است. اما اگر بخواهیم برای دانشجویان ارشد، ترمی را به‌عنوان پرخطر معرفی کرده باید دانشجوی دوره آموزشی پژوهشی رسمی با ۱ ترم را معرفی کرد. احتمال انصراف دانشجوی زن بیش از مرد است.

یافته‌های حاصل از تحلیل ارتباط مقطع، دوره تحصیلی و تعداد مشروطی و رتبه‌بندی آنها نشان داد که پرخطرترین دانشجویان با تعداد مشروطی عبارت‌اند از کارشناسی رسمی با ۷ یا ۸ یا ۹ ترم مشروطی، کارشناسی فراگیر با ۴ ترم مشروطی، ارشد آموزشی پژوهشی فراگیر با ۳ ترم مشروطی و دکتری رسمی بدون مشروطی جزو پرخطرترین‌ها هستند.

با توجه به آنکه نرخ انصراف دانشجویان ارشد دوره آموزشی پژوهشی رسمی ترم دوم ۵۶ درصد، ارشد آموزشی پژوهشی فراگیر و آموزشی فراگیر ترم سوم به ترتیب ۵۲ و ۴۱ درصد و کارشناسی دوره فراگیر ترم چهارم و پنجم ۴۴ درصد است باید نگاه ویژه به آنها داشته باشد.

هشدارهایی که قابل تعمق هستند:

- \* انصراف ۱۱ درصدی دانشجویان انصرافی ارشد زن ترم اول دوره آموزشی رسمی
- \* انصراف ۱۶ درصدی دانشجویان انصرافی ارشد مرد ترم اول دوره آموزشی رسمی

درصد) که در رأس آنها دانشجویان ۲۹ سال قرار دارند. فرضیه دوم تحقیق درخصوص نقش مقطع تحصیلی به‌تنهایی تأیید نمی‌شود اثر آن در صورتی تأیید می‌شود که این صفت با سن، دوره و گروه تحصیلی همراه باشد.

فرضیه سوم تحقیق (نقش گروه‌های تحصیلی) رد می‌شود زیرا تفاوت در گروه‌های مختلف تفاوت محسوسی در انصراف ایجاد نمی‌کند.

قوانین استنتاج بر مبنای الگوریتم پاسخ خود یادگیرنده به شرح زیر می‌باشند:

- ۴۷ درصد دانشجویان انصرافی با ۱۹ و ۵۰ درصد با ۲۰ سال سن، در ترم سوم انصراف می‌دهند.

- ۲۰ درصد دانشجویان انصرافی با ۲۱ سال سن، در ترم پنجم و ۱۴ درصد در ترم ششم انصراف می‌دهند.

یافته‌های قابل توجه الگوریتم پاسخ خود یادگیرنده درخصوص دوره تحصیلی به شرح زیر می‌باشند:

- ۵۶ درصد دانشجویان انصرافی ارشد دوره آموزشی پژوهشی رسمی در ترم دوم انصراف می‌دهند.

- ۵۲ درصد دانشجویان انصرافی ارشد دوره آموزشی پژوهشی فراگیر و ۴۱ درصد دانشجویان انصرافی ارشد دوره آموزشی فراگیر در ترم سوم انصراف می‌دهند.

- ۲۲ درصد دانشجویان انصرافی کارشناسی دوره فراگیر ترم چهارم و ۲۰ درصد در ترم پنجم انصراف می‌دهند.

یافته‌های قابل توجه الگوریتم پاسخ خود یادگیرنده در ارتباط با دوره و تعداد مشروطی به شرح زیر است:

- ۴۴ درصد دانشجویان انصرافی ارشد دوره آموزشی رسمی با ۰ یا ۱ مشروطی، ترم اول انصراف می‌دهند.

- ۳۷ درصد دانشجویان انصرافی کارشناسی بدون مشروطی، ترم دوم انصراف می‌دهند.

یافته‌های قابل توجه الگوریتم پاسخ خود یادگیرنده درخصوص تعداد مشروطی به شرح زیر می‌باشند:

- ۲۶ درصد دانشجویان انصرافی بدون مشروطی و ۵۰ درصد دانشجویان زن انصرافی بدون مشروطی در ترم دوم انصراف می‌دهند.

- ۳۱ درصد دانشجویان انصرافی با ۲ یا ۳ مشروطی در ترم چهارم انصراف می‌دهند.

- ۴۹ درصد دانشجویان انصرافی با بیش از ۳ مشروطی در ترم ۷، ۸ و ۹ انصراف می‌دهند.

- ۴۹ درصد دانشجویان انصرافی با ۳ یا بیش از ۳ مشروطی در ترم ششم و ۱۸ درصد در ترم ۹ و ۱۰ انصراف می‌دهند.

روی گروه‌های پر خطر تمرکز داشته باشد و برای کسب اطمینان از سیاست‌های اتخاذ شده آتی، هر سال با تحلیل اطلاعات جدید، تغییرات دسته‌های آسیب‌پذیر و نرخ انصراف را بررسی و آنگاه سیاست‌های بازدارنده، تشویقی و حمایتی متناسب را اتخاذ و اجرا کند. انجام تحقیقات مشابه برای استان‌های پرریزش نیز می‌تواند از اقدامات آتی دانشگاه پیام نور باشد.

\* انصراف ۳۵ درصدی دانشجویان انصرافی ارشد زن ترم چهارم دوره فراگیر الکترونیکی  
\* انصراف ۳۱ درصدی دانشجویان انصرافی ارشد مرد ترم چهارم دوره آموزشی پژوهشی فراگیر  
\* درخصوص سؤال چهارم و کمک نتایج تحقیق به مدیران ارشد دانشگاه می‌توان بیان داشت دانشگاه پیام نور استان باید بر

### منابع

کلانتری، خلیل (۱۳۹۱). پردازش و تحلیل داده‌ها در تحقیقات اجتماعی \_ اقتصادی با استفاده از نرم‌افزار SPSS. تهران: فرهنگ صبا، ۳۴۴-۳۲۹.

شهرابی، جمال (۱۳۹۰). *داده‌کاوی*. تهران: جهاد دانشگاهی، واحد دانشگاه صنعتی امیر کبیر، ۷۱-۲۴.  
شهرابی، جمال (۱۳۹۲). *داده‌کاوی با کلمنتاین*. تهران: جهاد دانشگاهی، واحد دانشگاه صنعتی امیر کبیر، ۱۷۰-۲۶۸.

Baker, R. & Yosef, K. (2009). "The state of educational data mining in 2009: A review and future visions". *Journal of Educational Data Mining*, 1(1), 3-17.

9890.  
Ho Yu, C., DiGangi, S., Jannasch-Pennell, A. & Kaprolet, Ch. (2010). "A Data Mining Approach for Identifying Predictors of Student from sophomore to Junior Year". *Journal of Data Science*, 8(2), 307-325.

Bienkowski M., Feng M. & Means B. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief*. tech. rep, U.S. Department of Education, Office of Educational Technology, Center for Technology in Learning, SRI International, 1, 1-57.

Sanjeev A. P. & Zytow J. M. (1995), "Discovering Enrollment Knowledge in University. Databases Regularities (1995)," *Data mining approach to student retention*". 1st International Conference on Knowledge Discovery and Data Mining (KDD-95), 246-251.

Bogard, M. & James, C. (2012). *Using SAS Enterprise BI and SAS Enterprise Miner TM to Reduce Student Attrition*. tech. rep, Paper presented at the 2012 SAS Global Forum, Orlando, FL. 1-10.

Dekker, G., Pechenizkiy, M. & Vleeshouwers, J. (2009). "Predicting students drop out: a case study". In 2nd International Educational Data Mining Conference, 2, 41-50.

Bogard, M. (2013). *A Data Driven Analytic Strategy for Increasing Yield and Retention at Western Kentucky University Using SAS Enterprise BI and SAS Enterprise Miner*. tech. rep, SAS Global Forum 2013 Proceedings. Cary, NC: SAS® Institute, 1-12 .

Druzdel, M. & Glymour, C. (2009). *Application of the TETRAD II program to the study of student retention in US colleges*. in AAAI-94 Workshop on Knowledge Discovery in Databases, KDD94, 419-430.

Campbell J. P. & Oblinger D. G. (2007), "Academic Analytics A New Tool for a New Era". *Educause Review*, 42(4), 40-57.

Laura, E., Baron, J. D., Devireddy, M. & Sundararaju, V. (2012). *Mining Academic Data to Improve College Student Retention: An Open Source Perspective*. In International Conference on Learning Analytics and Knowledge, 12, 139-142 .

Chin, K.S., Tang, D.W., Yang, J. B. & Wong, S. Y. & Wang H. (2009). Assessing new product development project risk by Bayesian network with a systematic probability generation methodology. *Expert Systems with Applications journal*, 36(6), 9879-

Emmet, C. & Mark, A. (2013). *Leading on the Edge of Chaos*, Prentice Hall Press; 1st edition.  
García, E. & Romero, C. Ventura S. & de

- Castro, C. (2011). "A collaborative educational association rule mining tool". *The Internet and Higher Education Journal*, 14(2), 77-88.
- Grebennikov, L. & Shah, M. (2012). "Investigating attrition trends in order to improve student retention". *Quality Assurance in Education*, 20(3), 223-236.
- Han, J. & Kamber, M. (2006). *Data Mining Concepts and Technique*. Elsevier, Third Edition, 7-40.
- Herzog, S. (2005). "Measuring Determinants of Student Return VS. Dropout/Stopout VS. Transfer: A First-to-Second Year Analysis of New Freshmen". *Research in Higher Education*, 46, 883-928.
- Herzog, S. (2010). "Estimating student retention and degree completion time: Decision trees and neural networks Vis-à-Vis regression". *New Directions for Institutional Research*, 131, 17-33.
- Huebner, R. A. (2012). "A Survey of Education Data Mining Research". *Research in Higher Education Journal*, 1-13.
- Jiawei Han and Micheline Kamber (2006). *Data Mining Concepts and Techniques*. Elsevier, Second Edition, 7-9.
- Lykourantzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis G. & Loumos, V. (2009). "Dropout prediction in e-learning courses through the combination of machine learning techniques". *Computers & Education*, 53, 950-965.
- Mello, S. K. D., Calvo, R. A. & Eds, A. O. (2013). "student retention, case study". Proceedings of the 6th International Conference on Educational Data Mining, 316-317.
- Mohammed Zafaruddin, G. H. & Jadhav, H. (2013). "Data mining approach to student retention". *International Monthly Refereed Journal of Research In Management & Technology*, 2, 98-102
- Nandeshwar A. and Chaudhari S. (2009), "Enrollment prediction models using data mining". Retrieved January journal, 2, 1-17.
- Owens, J. (2009). "Student Withdrawal from Higher Education information". *Tech. Rep DCELLS Welsh Assembly Government*, 22-63.
- Pittman, K. (2008). *Comparison of data mining techniques used to predict student retention*. PhD thesis, Nova Southeastern University, (UMI No. 3297573, 48-73.
- Romero, C. & Ventura, S. (2013). "Data mining in education". *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27.
- Scott, G., Shah, M., Grebennikov, L. & Singh, H. (2008). "Improving student retention: A University of Western Sydney case study". *Journal of Institutional Research*, 14(1), 9-23.
- Superby, J., Vandamme, J. & Meskens, N. (2006). "Determination of factors influencing the achievement of the first-year university students using data mining methods". In *Workshop on Educational Data Mining*, 37-44.
- Zhang, Y., Oussena, S., Clark, T. & Kim H. (2010). *Use Data Mining To Improve Student Retention in Higher Education - A Case Study*. in 12th International Conference on Enterprise Information Systems (ICEIS), 190-197.
- Zlatko (2010). *Early prediction of student success: Mining students enrolment data*. in Informing Science & IT Education Conference (In SITE), 647-665.
- Scheuer, O. & McLaren, B. M. (2011). "Educational Data Mining". In *the Encyclopedia of the Sciences of Learning*, Springer, 1075-1079.
- Thearling K. (2013),. "An Introduction to Data Mining". *Research In Management & Technology journal*, 39(2), 105.
- Zaiane, O. R.(2013). *Introduction to Data Mining. CMPUT690 Principles of Knowledge Discovery in Databases*. Department of Computer Science, University of Alberta, 1-15.

