

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



بسمه تعالی

تاییدیه اعضای هیات داوران حاضر در جلسه دفاع از پایان نامه

آقای بهمن امینی پایان نامه ۶ واحدی خود را با عنوان اثر تغییرات دینامیکی پارامتر
زمان در محاسبه احتمال رویگردانی مشتری در مدلسازی ارزش عمر مشتری در
تاریخ ۱۳۹۱/۸/۱۵ ارائه کردند.

اعضای هیات داوران نسخه نهایی این پایان نامه را از نظر فرم و محتوا تایید کرده و
پذیرش آنرا برای تکمیل درجه کارشناسی ارشد مهندسی فناوری اطلاعات - سیستمهای
اطلاعاتی پیشنهاد می کنند.

عضو هیات داوران	نام و نام خانوادگی	رتبه علمی	امضا
استاد راهنما	دکتر امیر البدوی	استاد	
استاد مشاور	دکتر بابک تیمورپور	استادیار	
استاد ناظر	دکتر محمد مهدی سپهری	دانشیار	
استاد ناظر	دکتر محمدجعفر تارخ	دانشیار	
مدیر گروه (یا نماینده گروه تخصصی)	دکتر محمد مهدی سپهری	دانشیار	

آیین نامه چاپ پایان نامه (رساله) های دانشجویان دانشگاه تربیت مدرس

نظر به اینکه چاپ و انتشار پایان نامه (رساله) های تحصیلی دانشجویان دانشگاه تربیت مدرس، مبین بخشی از فعالیتهای علمی - پژوهشی دانشگاه است بنابراین به منظور آگاهی و رعایت حقوق دانشگاه، دانش آموختگان این دانشگاه نسبت به رعایت موارد ذیل متعهد می شوند:

ماده ۱: در صورت اقدام به چاپ پایان نامه (رساله) ی خود، مراتب را قبلاً به طور کتبی به «دفتر نشر آثار علمی» دانشگاه اطلاع دهد.

ماده ۲: در صفحه سوم کتاب (پس از برگ شناسنامه) عبارت ذیل را چاپ کند:

«کتاب حاضر، حاصل پایان نامه کارشناسی ارشد/ رساله دکتری نگارنده در رشته مهندسی فناوری اطلاعات است که در سال

۱۳۹۱ در دانشکده فنی و مهندسی دانشگاه تربیت مدرس به راهنمایی سرکار

خانم/جناب آقای دکتر امیرالدینی، مشاوره سرکار خانم/جناب آقای دکتر بابک تیمورپور

و مشاوره سرکار خانم/جناب آقای دکتر از آن دفاع شده است.»

ماده ۳: به منظور جبران بخشی از هزینه های انتشارات دانشگاه، تعداد یک درصد شمارگان کتاب (در هر نوبت چاپ) را به «دفتر نشر آثار علمی» دانشگاه اهدا کند. دانشگاه می تواند مازاد نیاز خود را به نفع مرکز نشر در معرض فروش قرار دهد.

ماده ۴: در صورت عدم رعایت ماده ۳، ۵۰٪ بهای شمارگان چاپ شده را به عنوان خسارت به دانشگاه تربیت مدرس، تأدیه کند.

ماده ۵: دانشجو تعهد و قبول می کند در صورت خودداری از پرداخت بهای خسارت، دانشگاه می تواند خسارت مذکور را از طریق مراجع قضایی مطالبه و وصول کند؛ به علاوه به دانشگاه حق می دهد به منظور استیفای حقوق خود، از طریق دادگاه، معادل وجه مذکور در ماده ۴ را از محل توقیف کتابهای عرضه شده نگارنده برای فروش، تامین نماید.

ماده ۶: اینجانب بهمن امینی دانشجوی رشته مهندسی فناوری اطلاعات مقطع کارشناسی ارشد

تعهد فوق و ضمانت اجرایی آن را قبول کرده، به آن ملتزم می شوم.

نام و نام خانوادگی: بهمن امینی

تاریخ و امضا:


۹۲۱/۳۱



دانشگاه تربیت مدرس

دانشکده فنی و مهندسی

بخش صنایع - گروه فناوری اطلاعات

پایان نامه برای دریافت درجه کارشناسی ارشد رشته مهندسی فناوری اطلاعات

گرایش سیستم های تکنولوژی اطلاعات

اثر تغییرات دینامیکی پارامتر زمان بر محاسبه احتمال رویگردانی

مشتری در مدلسازی ارزش عمر مشتری

بهمن امینی

استاد راهنما:

دکتر امیر البدوی

استاد مشاور:

دکتر بابک تیمورپور

مهر ۱۳۹۱

پاس بی کران بر آفریننده دانش، خردواندیش

تقدیم بہ:

پدر و مادرم بہ خاطر زحمات بی دریغشان

حیدم کللی زباغ ادب تابرور عمید

دربارگاہ میر ادب پرور آورم

حیف است باخسان گل دانش کنی نثار

من گل نثار مردم دانشور آورم

شہریار

در پشت بعضی اسم ها و کلمه های توانی سکر بگیری

می توانی به احترام بعضی اسم ها بلند شوی، بایستی و با افتخار از آنها یاد کنی

می توانی در مقابل عظمت اراده و از خود گذشته گشتی شان تعظیم کنی، مهوت بانی و در مقابل شان تنها شکر کنی...

تقدیر و شکر بسیار از استاد فریخته، فرزانه و کرانه جناب آقای دکتر البدوی و جناب آقای دکتر

تیمور پور که همواره با همراهی های بی دریغ و میان نکته های راه کشا، راهنما و آموزگار نگارنده در اتمام تحقیق

حاضر بودند

از شرکت ارتباطات یار ایران (همراه اول) به ویژه جناب آقای دکتر سید حسن سید موسوی (مدیر کل دفتر

تحقیقات و مطالعات راهبردی) و جناب آقای مهندس صحرانور (سرپرست معاونت فناوری اطلاعات) به

دلیل همکاری های انجام گرفته در این پژوهش نهایت تشکر و قدردانی را دارم.

این پژوهش با حمایت موسسه تحقیقات ارتباطات و فناوری اطلاعات ایران به شماره ۱۹۲۳۰/۵۰۰/ت

انجام شده است.

چکیده

امروزه در محیط رقابت کسب و کار، پدیده رویگردانی مشتری یک نگرانی عمده برای اکثر سازمان‌هایی محسوب می‌شود که در صنایعی با هزینه جابجایی اندک و هزینه بالای کسب مشتری فعالیت دارند. در این بین صنعت تلفن همراه با نرخ تقریبی رویگردانی ۳۰٪ در سال، از جمله مطرح‌ترین صنایع محسوب می‌شود. محققان برای غلبه بر این پدیده رویکردهای مختلفی از جمله توسعه مدل برای پیش‌بینی رویگردانی ارائه داده‌اند. در اکثر تحقیقات به منظور مدل‌سازی رویگردانی در صنعت تلفن همراه، از متغیرهای رفتاری مربوط به مشتریان به منظور پیش‌بینی استفاده شده است. این مطالعات با این فرض انجام می‌گیرند که این متغیرهای رفتاری، متغیرهای مستقلی هستند که می‌توان به کمک آن‌ها متغیر وابسته‌ای همچون رویگردانی مشتری را با روش‌های گوناگون داده‌کاوی محاسبه نمود. در واقع آن‌ها به طور تلویحی این فرض را دارند که متغیر دیگری بر رفتار مشتری تأثیری ندارد و این متغیرهای رفتاری، واقعیت رفتاری هر مشتری را در تعامل با صنعت مربوطه نشان می‌دهند. در این تحقیق ما از پارامتر زمان به عنوان یک متغیر دینامیکی مستقل که در برهه‌هایی از زمان بر برخی از متغیرهای مستقل همچون تراکنش‌های مشتری، اثر گذار است بهره خواهیم برد تا بتوانیم رویگردانی مشتری را با دقت بهتری پیش‌بینی نماییم. بدین منظور یک متدولوژی سه مرحله‌ای که شامل کشف نقاط زمانی اثرگذار، کشف مشتریان تحت تأثیر قرار گرفته و تعدیل متغیرهای رفتاری هر یک از آن‌ها است، ارائه شده است. نتایج این تحقیق حاکی از آن است که، به طور کل پارامتر زمان به طور دینامیکی تأثیر شگرفی بر روی متغیرهای رفتاری مشتری دارد. همچنین پس از لحاظ نمودن اثر این پارامتر بر روی تراکنش‌های مشتریان، دقت پیش‌بینی متغیر رویگردانی با استفاده از معیار فایده (۰.۱۰) به طور میانگین ۲٪ بهبود داشته است.

واژگان کلیدی: رویگردانی مشتری، اثر دینامیکی پارامتر زمان، آزمون لیون، داده کاوی،

تعدیل، صنعت تلفن همراه، تنظیم قراردادی

فهرست مطالب

۱	فصل اول: معرفی و کلیات
۱-۱	مقدمه
۲-۱	مسئله تحقیق
۳-۱	هدف از انجام تحقیق
۴-۱	سؤالات تحقیق
۵-۱	اهمیت و ضرورت انجام تحقیق
۶-۱	مفروضات تحقیق
۷-۱	تعاریف و اصطلاحات عمده
۸-۱	مشارکت تحقیق
۱-۸-۱	مشارکت تئوریک
۲-۸-۱	مشارکت عملی
۹-۱	ساختار گزارش پایان نامه
۹-۱	نتیجه گیری
۹	فصل دوم: مرور ادبیات
۱-۲	مقدمه
۲-۲	مدیریت ارتباط با مشتری: مفاهیم پایه
۳-۲	داده کاوی و کاربرد های آن در مدیریت ارتباط با مشتری
۱-۳-۲	دسته بندی
۴-۲	رویگردانی مشتری: مرور ادبیات
۱-۴-۲	تعریف رویگردانی مشتری
۲-۴-۲	انواع رویگردانی مشتری
۳-۴-۲	بررسی و تحلیل عوامل رویگردانی مشتری
۴-۴-۲	مدل سازی رویگردانی مشتری
۵-۲	داده کاوی در حالت کمیایی
۶-۲	نقد ادبیات
۷-۲	نتیجه گیری
۶۹	فصل سوم: روش شناسی و داده های تحقیق
۱-۳	مقدمه

۷۰	۲-۳ طراحی تحقیق.....
۷۰	۱-۲-۳ هدف تحقیق.....
۷۱	۲-۲-۳ رویکرد تحقیق.....
۷۲	۳-۲-۳ استراتژی تحقیق.....
۷۴	۳-۳ فرایند تحقیق.....
۷۵	۱-۳-۳ جمع آوری داده.....
۷۵	۲-۳-۳ انتخاب داده.....
۷۶	۳-۳-۳ پیش پردازش داده.....
۸۷	۴-۳-۳ تبدیل داده.....
۸۷	۵-۳-۳ داده کاوی.....
۸۷	۶-۳-۳ ارزیابی / تفسیر.....
۹۲	۴-۳ نتیجه گیری.....
۹۳	فصل چهارم : تحلیل داده های تحقیق.....
۹۴	۱-۴ مقدمه.....
۹۴	۲-۴ نرم افزار های مورد استفاده در تحقیق.....
۹۵	۳-۴ تعدیل تراکنشات مشتریان.....
۹۵	۱-۳-۴ کشف مشتریان متأثر از زمان.....
۱۰۰	۲-۳-۴ تعدیل تراکنشات مشتریان متأثر از زمان.....
۱۰۷	۴-۴ تبدیل داده.....
۱۰۹	۵-۴ ساخت مدل پیش بینی کننده.....
۱۱۷	۶-۴ نتیجه گیری.....
۱۱۸	فصل پنجم: نتیجه گیری.....
۱۱۹	۱-۵ مقدمه.....
۱۱۹	۲-۵ مروری بر فصل های گذشته.....
۱۲۰	۳-۵ پاسخ سوالات تحقیق.....
۱۲۱	۴-۵ نوآوری های پژوهش.....
۱۲۱	۱-۴-۵ نوآوری در سطح تئوری.....
۱۲۲	۱-۴-۵ نوآوری در سطح بکارگیری.....
۱۲۲	۵-۵ محدودیت های پژوهش.....
۱۲۳	۶-۵ پژوهش های آتی.....

- ۱۲۴ نتیجه گیری ۷-۵
- ۱۲۵ مراجع
- ۱۲۹ پیوست ا: آزمون های آماری
- ۱۳۳ پیوست ب: الگوریتم های درخت تصمیم
- ۱۳۷ پیوست پ: کدهای مورد استفاده در تحقیق
- ۱۴۰ واژه نامه انگلیسی به فارسی

فهرست جداول

- جدول ۱-۲: دوره مشاهده، حفظ و پیش بینی برای هر خوشه از مشتریان (Jahromi et al., 2010)..... ۶۰
- جدول ۲-۲: عملکرد مدل پیش بینی کننده بر اساس معیار فایده (Jahromi et al., 2010)..... ۶۱
- جدول ۳-۲: عملکرد مدل پیش بینی کننده حساس به هزینه بر اساس معیار فایده (Jahromi et al., 2010)..... ۶۲
- جدول ۱-۳: تحقیق کمی و تحقیق کیفی (Malhotra, 2007)..... ۷۲
- جدول ۲-۳: ماتریس مصرف در روزهای متوالی برای مشتریان متأثر از زمان..... ۸۲
- جدول ۱-۴: عملکرد مدل پیش بینی کننده C5.0 برای سه مجموعه داده با استفاده از معیار فایده..... ۱۱۰
- جدول ۲-۴: میزان صحت و مقدار AUC مدل ساخته شده به کمک الگوریتم C5.0 بر روی هر سه مجموعه داده..... ۱۱۳
- جدول ۳-۴: عملکرد مدل پیش بینی کننده CART برای سه مجموعه داده با استفاده از معیار فایده..... ۱۱۴
- جدول ۴-۴: عملکرد مدل پیش بینی کننده CHAID برای سه مجموعه داده با استفاده از معیار فایده..... ۱۱۴
- جدول ۵-۴: عملکرد مدل پیش بینی کننده ANN برای سه مجموعه داده با استفاده از معیار فایده..... ۱۱۴
- جدول ۶-۴: میزان صحت و مقدار AUC مدل ساخته شده به کمک الگوریتم CART..... ۱۱۵
- جدول ۷-۴: میزان صحت و مقدار AUC مدل ساخته شده به کمک الگوریتم CHIAD..... ۱۱۶
- جدول ۸-۴: میزان صحت و مقدار AUC مدل ساخته شده به کمک روش ANN..... ۱۱۶

فهرست اشکال

- شکل ۱-۲: نمایش چرخه عمر مشتری (Olafsson et al., 2008) ۱۲
- شکل ۲-۲: رشد تولید مقالات در حوزه CRM (منبع: پایگاه داده SSCI) (Hsu-Hao Tsai, 2012) ۱۶
- شکل ۳-۲: روند تکامل جستجو برای کسب اطلاعات (Lejeune, 2001) ۱۷
- شکل ۴-۲: واحد یک شبکه عصبی مصنوعی که از روی یک نرون بیولوژیکی مدل سازی شده است. خروجی واحد ترکیب غیر خطی از ورودی‌هایش است (Berry and Linoff, 1997) ۲۲
- شکل ۵-۲: مدلی مفهومی برای رویگردانی مشتری با اثرات واسطه‌ای (Ahn et al., 2006) ۳۰
- شکل ۶-۲: مدل مفهومی رفتار حفظ مشتری در سرویس بی سیم (Seo et al., 2008) ۳۳
- شکل ۷-۲: زمان‌بندی منطقی ثبت مقادیر متغیرها در طول یک سال مشاهده (Ahn et al., 2006) ۳۶
- شکل ۸-۲: نتایج آزمایش - اثرات مطلوب نسبت‌های کلاس (استفاده از ۲ زیر دوره) (Wei and Chiu, 2002) ۴۴
- شکل ۹-۲: چهار چوبی برای پیش بینی رویگردانی مشتری و حداکثر سازی سودآوری (Mozer et al., 2000) ۴۹
- شکل ۱۰-۲: عملکرد آزمون برای چهار روش پیش بینی کننده رویگردانی مشتری. پیش بینی کننده‌ها بر اساس مدل‌های مختلف (رگرسیون لجستیک، شبکه عصبی) و نمایش‌های ورودی (ساده، پیچیده) هستند (Mozer et al., 2000) ۵۱
- شکل ۱۱-۲: عملکرد آزمون برای چهار روش پیش بینی کننده رویگردانی مشتری. پیش بینی کننده‌ها بر اساس مدل‌های مختلف (درخت تصمیم، شبکه عصبی) و اینکه نمونه‌گیری انجام شود یا نشود است (Mozer et al., 2000) ۵۱
- شکل ۱۲-۲: نمای کلی معماری CHAMP (Datta et al., 2000) ۵۳
- شکل ۱۳-۲: میزان لیفت برای $n\%$ بالا (۵٪ تا ۱۰۰٪) (Datta et al., 2000) ۵۷
- شکل ۱۴-۲: بازدهی تجمعی (در هزار) برای $n\%$ بالا (۵٪ تا ۱۰۰٪) (Datta et al., 2000) ۵۷
- شکل ۱۵-۲: میزان تغییرات لیفت برای دهک اول مشتریان در گذر زمان (Datta et al., 2000) ۵۸
- شکل ۱۶-۲: منحنی لیفت مربوط به الگوریتم‌های مختلف رندوم فرست (Xie et al., 2009) ۶۶
- شکل ۱۷-۲: منحنی لیفت مربوط به الگوریتم‌های مختلف (Xie et al., 2009) ۶۶
- شکل ۱-۳: فلوجارت کشف دانش در پایگاه داده (Li and Ruan, 2007) ۷۴
- شکل ۲-۳: میزان پاسخ تجمعی به ارسال ایمیل هدف گذاری شده در مقایسه با ارسال ایمیل انبوه (Berry and Linoff, 1997) ۸۹

- شکل ۳-۳: ماتریس درهم ریختگی (He and Garcia, 2009)..... ۹۰
- شکل ۱-۴: زمانبندی منطقی داده‌ها جهت ساخت مدل طی مدت ۲ ماه..... ۱۰۸
- شکل ۲-۴: سه دوره زمانی با اثرگذاری متفاوت پارامتر زمان..... ۱۰۸
- شکل ۳-۴: نمودار فایده الگوریتم **C5.0** برای مجموعه داده **WOTEA**..... ۱۱۱
- شکل ۴-۴: نمودار فایده مربوط به الگوریتم **C5.0** برای مجموعه داده **WTEAMD**..... ۱۱۱
- شکل ۵-۴: نمودار فایده مربوط به الگوریتم **C5.0** برای مجموعه داده **WTEAIR**..... ۱۱۲

فهرست نمودارها

- نمودار ۱-۳: میانگین روزانه مدت زمان تماس به ثابیه برای تمام مشتریان ۷۷
- نمودار ۲-۳: فرایند تعدیل مقادیر مصرفی مشتریان در مرحله پیش پردازش ۸۶
- نمودار ۱-۴: نمودار پراکندگی میانگین تماس مشتریان در روزهای غیر تعطیل و تعطیل ۹۶
- نمودار ۲-۴: نمودار پراکندگی میانگین فراوانی تماس مشتریان در روزهای غیر تعطیل و تعطیل ۹۷
- نمودار ۳-۴: سری زمانی میانگین مدت زمان تماس به ثابیه پس از تعدیل مدت زمان تماس مشتریان متأثر از زمان به روش
تفاضل میانگین ۱۰۲
- نمودار ۴-۴: سری زمانی میانگین فراوانی تماس پیش از تعدیل فراوانی تماس مشتریان متأثر از زمان ۱۰۳
- نمودار ۵-۴: سری زمانی میانگین فراوانی تماس پس از تعدیل فراوانی تماس مشتریان متأثر از زمان به روش تفاضل میانگین ۱۰۴
- نمودار ۶-۴: سری زمانی میانگین مدت زمان تماس به ثابیه پس از تعدیل مدت زمان تماس مشتریان متأثر از زمان به روش
معکوس رگرسیون ۱۰۵
- نمودار ۷-۴: نمودار پراکندگی میانگین زمان تماس مشتریان در روزهای غیر تعطیل و تعطیل ۱۰۶

فصل اول: معرّفی و کلیات

۱-۱- مقدمه

امروزه جذب و حفظ مشتریان به عنوان یکی از دغدغه های عمده رهبران کسب و کار محسوب می شود. در حالی که شرکت های نوظهور بر روی جذب مشتریان جدید تمرکز دارند، شرکت های به بلوغ رسیده جهت ایجاد فرصت های خرید جانبی برای مشتریان و افزایش ارزش طول عمر آنها تمایل به حفظ مشتریان موجود دارند.

در عصر جدید، ظهور تجارت الکترونیک دسترسی به اطلاعات را به طور فزاینده ای ارتقا داده است. پیارد (۲۰۰۰) بر این باور است که اینترنت، قدرت تصمیم گیری مشتریان ارتقا داده بطوریکه تا پیش از این تنها یک انتخاب به عنوان تأمین کننده داشتند. در حالی که اکنون رقیبان به اندازه یک کلیک با مشتریان فاصله دارند. در نتیجه افزایش قدرت تصمیم گیری مشتریان می تواند منجر به افزایش نرخ رویگردانی سازمان ها شود (Lejeune, 2001). برای مقابله با چنین تهدیداتی، شرکت ها بایستی به روش های کارا و موثری مجهز باشند تا رفتارهای احتمالی مشتریان را پیش از ترک پیش بینی نمایند.

هدف مطالعه حاضر، ارائه مفهومی نو در ادبیات رویگردانی و در نتیجه مدل سازی دقیق تر رویگردانی مشتری در صنعت تلفن همراه به کمک روش های یادگیری ماشین است.

۱-۲- مسئله تحقیق

در صنعت تلفن همراه نیز تحقیقات زیادی به منظور مدل سازی یا پیش بینی رویگردانی مشتری انجام گرفته است که هر کدام بنا به میزان تعهد طرفین (قراردادی^۱ یا غیر قراردادی^۲)، میزان

^۱ Contractual setting

^۲ Non contractual setting

دسترسی به متغیرهای رفتاری مشتری، ویژگی‌های جمعیت شناختی هر یک و نیز مشکلات مدل سازی از روش‌ها و استراتژی‌های گوناگونی بهره برده‌اند. در اکثر این تحقیقات، محقق با استفاده از این متغیرهای رفتاری مستقل و ویژگی‌های مشتریان و نوع صنعت، سعی در پیش بینی متغیر وابسته رویگردانی، برای هر یک از مشتریان است. در واقع تاکنون محققین تلویحاً فرض بر مستقل بودن متغیرهای رفتاری مشتریان داشته‌اند و مقادیر این متغیرها را ناشی از رفتار معمول مشتریان در بازه زمانی مورد مطالعه دانسته‌اند. در این پژوهش سعی شده است که ابتدا میزان اثر پارامتر زمان به عنوان عاملی مستقل که به طور دینامیکی متغیرهای رفتاری اکثر مشتریان (صرف نظر از تغییر وفاداری واقعی او) را در بازه های خاصی از زمان به طور معناداری متأثر می‌سازد، شناخته و سپس این تغییر دینامیکی و غیر معمول بر روی متغیرهای مصرفی را با روش‌های پیشنهادی حذف نموده تا بتوانیم متغیرهای رفتاری مشتری را مستقل‌تر نموده و در نتیجه پیش بینی رویگردانی را با دقت بهتری نسبت به گذشته، به کمک روش‌های داده کاوی، مدل سازی نماییم.

۱-۳- هدف از انجام تحقیق

هدف این تحقیق ایجاد و توسعه مدل رویگردانی مشتری با در نظر گیری اثر تغییرات دینامیکی پارامتر زمان بر پیش بینی رویگردانی مشتری در صنعت تلفن همراه (بازار تلفن همراه در حالت پس پرداخت^۳) است.

³ Post paid