



بهبود دقت تشخیص حساب‌های جعلی در توییتر (ایکس) با استفاده از ویژگی‌های محتوایی و محاسباتی در توییت‌ها و پروفایل‌ها

علیرضا نیکیان^۱، لیلا سلطانی تهرانی^۲

^۱ گروه مهندسی کامپیوتر، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران؛ nikian_a@iau.ac.ir

^۲ گروه مهندسی کامپیوتر، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران؛ leyla.s6750@gmail.com

* نویسنده مسئول: علیرضا نیکیان

چکیده:

هر روزه حجم بالایی از اطلاعات در شبکه‌های اجتماعی در حال انتشار است، این حجم عظیم از اطلاعات مشکلاتی را به همراه دارد. یکی از این مشکلات منتشر شدن محتوای غیرمجاز و مخرب است، این محتوا توسط حساب‌های جعلی تولید و منتشر می‌شود و باعث به وجود آمدن مشکلات و آسیب‌های جدی برای کاربران، نهادهای تجاری و دیگران در جامعه و دنیا خواهد شد. این پژوهش مشخص می‌کند که با ترکیب ویژگی‌های موجود و انجام محاسبات منطقی و عددی روی داده‌ها و ویژگی‌ها می‌توان به ویژگی‌های جدید و تاثیرگذار دست یافت. آزمایش‌های انجام شده در این پژوهش بر روی دو مجموعه داده‌ها صورت گرفته است که مجموعه داده‌هایی که بر اساس ویژگی‌های مبتنی بر پروفایل و توییت می‌باشد شامل ویژگی‌های جدید (مجموع هشتگ‌ها، ریتوییت‌ها، لایک‌ها، URL‌ها و منش‌ها در توییت‌ها و نسبت توییت‌های ریتوییت شده، لایک شده، دارای هشتگ، دارای URL و دارای منش‌ها به کل توییت‌ها) است و مجموعه داده‌هایی که دارای ویژگی‌های مبتنی بر پروفایل می‌باشد شامل ویژگی‌های جدید (لایک‌های بیشتر از ۱۰۰، طول نام کاربری کمتر از ۱۰ کاراکتر، موقعیت جغرافیایی فعال و همچنین نام کاربری غیر پیشفرض و دارای حداقل یک لیست، فیلد مکان پر و همچنین نام کاربری و تصویر پروفایل غیر پیشفرض، استفاده از بنر پروفایل کاربری و نسبت دنبال کنندگان به سن حساب) است که با استفاده از پنج الگوریتم دسته‌بندی انجام شده و در هر بخش نتایج پژوهش با دو کار انجام شده در گذشته مقایسه شده است و طبق نتایج به دست آمده، نشان داده شده است که با استفاده از ویژگی‌های پیشنهادی در این پژوهش، در الگوریتم درخت‌های ارتقا یافته گرایان دقت ۹۹/۲۶٪، الگوریتم درخت تصمیم دقت ۹۹/۱۹٪، الگوریتم جنگل تصادفی دقت ۹۹/۰۵٪، الگوریتم J48 دقت ۹۹/۳۴٪ و الگوریتم نزدیکترین همسایه دقت ۹۸/۸۷٪ به دست آمده است که نشان از بهبود دقت در اکثر نتایج به دست آمده دارد.

کلمات کلیدی: دقت تشخیص، حساب جعلی، توییتر، پروفایل، الگوریتم J48

۱- مقدمه

امروزه شبکه‌های اجتماعی ابزاری موثر در ایجاد ارتباط میان افراد شده است. کاربران در این شبکه‌ها با اشتراک گذاری علایق، نمونه کار، تجربیات، محصولات خود و دیدگاه‌ها با دیگران به تعامل می‌پردازند. در حال حاضر بیش از ۲ میلیارد نفر در شبکه‌های اجتماعی نظیر فیس‌بوک، توییتر، لینکدین و گوگل پلاس عضو هستند. شبکه اجتماعی توییتر یک میکرو بلاگ است که کاربران در آن می‌توانند مطالبی به طول ۲۸۰ کاراکتر ارسال کنند. این شبکه‌ی اجتماعی محبوب بیش از ۳۲۸ میلیون کاربر عضو دارد و ۱۵۷ میلیون کاربر فعال روزانه، که این کاربران روزانه بیش از ۵۰۰ میلیون توییت ارسال می‌کنند. آمار ارائه شده توسط انور و همکاران در سال ۲۰۱۶ برای حساب‌های جعلی حدود ۴۸ میلیون حساب است.

حساب‌های جعلی در توییتر از نظر نوع فعالیت :

حساب‌هایی به منظور ترویج یا تخریب محصول و یا یک سرویس خاص.

حساب‌هایی که به قصد سوء استفاده‌های مالی در این شبکه اجتماعی فعالیت می‌کنند.

حساب‌های منتشر کننده اخبار جعلی

حساب‌هایی با قصد آزار و اذیت دیگران

میزان قابل توجه درآمد در بازار سیاه خرید و فروش حساب‌های جعلی و پتانسیل به وجود آمدن آسیب‌های جدی برای کاربران توسط این حساب‌ها، از قبیل اذیت و آزار سایبری، کلاهبرداری‌های مالی و حملات سرقت هویت باعث شده که شناسایی و جلوگیری از فعالیت این حساب‌ها

یکی از الزامات مهم و پر اهمیت باشد.



تجارتی با درآمد سالانه ۴۰ تا ۳۶۰ میلیون دلار



۱۸ دلار به ازای هر ۱۰۰۰ حساب جعلی



۵٪ توپیتهای دارای URL مخرب هستند

شکل ۱: انواع حسابهای جعلی در توپیتر



شکل ۲: هدف مقاله

۲- شبکه های اجتماعی

سرویس های مبتنی بر وب که کاربران در آنها نظرات، علاقه مندی ها و اطلاعات خود را با یکدیگر به اشتراک می گذارند.

۲-۱- فرایند آماده سازی و پیش پردازش داده ها

- پاک سازی داده
- یکپارچه سازی داده
- تبدیل داده



شکل ۳: ساختار انتخاب ویژگی

۲-۲- معیارهای ارزیابی

- ماتریس درهم ریختگی
- ارائه تجسمی از کارایی یک مدل یادگیری با نظارت
- معیار صحت یا درستی

مهم ترین معیار برای تعیین کارایی یک الگوریتم دسته بندی

▪ معیار دقت

درجه تکرارپذیری نتایج، طی اندازه گیری های متوالی را Precision می نامند.

$$\text{Precision} = \frac{TP}{TP+FP}$$

▪ معیار بازخوانی

دقت تشخیص دسته بندی درست را محاسبه می کند

$$\text{Recall} = \frac{TP}{TP+FN}$$

▪ معیار F-measure

توصیف کننده میانگین وزن دار مابین دو کمیت Precision و Recall، برای یک الگوریتم کلاس بندی کننده در شرایط ایده آل، مقدار این کمیت برابر با ۱ می باشد.

$$\text{F-measure} = \frac{2 \times (\text{Recall} \times \text{Precision})}{\text{Recall} + \text{Precision}}$$

▪ منحنی ROC

در این منحنی نرخ تشخیص مثبت واقعی (TPR) روی محور Y و نرخ تشخیص مثبت کاذب (FPR) روی محور X رسم می شود. مساحت زیر این منحنی یا به اصطلاح AUC نشان دهنده کیفیت یک مدل دسته بندی می باشد

$$\text{TPR} = \frac{TP}{TP+FN}$$

$$\text{FPR} = \frac{TN}{TN+FP}$$

۳- پیشنهاد پژوهش

به طور کلی پژوهش های انجام شده در این زمینه با یکی از ۳ نگرش زیر به حل مساله پرداخته اند:

۱- مبتنی بر رفتار: بیشتر مدل های مبتنی بر رفتار بر اساس نظریه آماری یا الگوریتم خوشه بندی می باشند.

۲- مبتنی بر گراف: الگوریتم های مبتنی بر گراف بر روابط میان حساب های کاربری در شبکه های اجتماعی آنلاین متمرکز می باشند.

۳- مبتنی بر دسته بندی و یادگیری ماشین: اکثر روش های یادگیری ماشین با استفاده از الگوریتم های یادگیری ماشینی به آموزش دسته بندی می پردازند.

جدول ۱: پژوهش های مبتنی بر رفتار

سال	پژوهش گر	شبکه اجتماعی	توضیحات
۲۰۱۳	وانگ و همکاران	RenRen, Linkedin	با بررسی ۱۶۰۰۰ جریان کلیک و خوشه بندی رفتارهای مختلف، موفق به ایجاد یک سیستم تشخیص سیبل شدند
۲۰۱۰	لین و همکاران	Facebook	بررسی لایک های فیس بوک، پیشنهاد یک سیستم تشخیص مبتنی بر الگوریتم های خوشه بندی برای جلوگیری از ترویج ناعادلانه لایک ها
۲۰۱۵	فو و همکاران	Twitter	شناسایی حساب های مخرب با استفاده از محاسبه دقت کاربران سرمایه داران اجتماعی نسبت به دیگر کاربران از دقت پایین تری برخوردارند

جدول ۲: پژوهش‌های مبتنی بر گراف

سال	پژوهش‌گر	شبکه اجتماعی	توضیحات
۲۰۱۲	گاش و همکاران	Twitter	آن‌ها "مزرعه پیوند" را به صورت یک حمله توصیف نمودند، معرفی رتبه تقلب که رویکردی مشابه به رتبه صفحه در مبارزه با مزرعه پیوند در تویتر دارد
۲۰۱۰	یاماگوچی و همکاران	Twitter	ارائه یک گراف کاربر-توییت، محاسبه اختیار کاربران، اکثر کاربران معتبر پیام‌ها را بدون بررسی دقیق منابع آن‌ها بازنشر یا ارسال می‌کنند.

جدول ۳: پژوهش‌های مبتنی بر دسته‌بندی و یادگیری ماشین

سال	پژوهش‌گر	شبکه اجتماعی	توضیحات
۲۰۱۰	الکس وانگ	Twitter	استفاده از مجموعه داده مشکل از حدود ۲۵ هزار داده‌ی کاربر، ۵۰۰ هزار توییت و ۴۹ میلیون روابط دوست/دنبال‌کننده، دریافت دقت ۸۹٪ با الگوریتم بی‌زین ساده
۲۰۱۰	بنونوتو و همکاران	Twitter	استفاده از ویژگی‌های محتوا و رفتاری، تشخیص ۷۰٪ از اسپم‌ها و ۹۶٪ از غیر اسپم‌ها با استفاده از الگوریتم SVM
۲۰۱۶	کلارک و همکاران	Twitter	استفاده از زبان طبیعی و مقایسه فراوانی ویژگی‌های متنی مانند نام، آدرس ایمیل و غیره به منظور طبقه‌بندی حساب‌های تویتر به عنوان مخرب یا قانونی

۴- روش پیشنهادی

در این پژوهش از رویکرد نظارتی یادگیری ماشین بر مبنای الگوریتم های دسته بندی استفاده شده است که الگوریتم مختلفی در این زمینه وجود دارند. در این تحقیق پس از تقسیم بندی مجموعه داده به عنوان مجموعه تست و آزمایش با استفاده از روش K fold cross validation، درخت تصمیم، جنگل تصادفی، درختهای ارتقا یافته گرادیان، k نزدیکترین همسایه و J48 استفاده شده است. نکته ای که این پژوهش را از نظر استفاده از الگوریتم های دسته بندی با پژوهش های دیگر متمایز میکند این میباشد که الگوریتم درختهای ارتقا یافته گرادیان یک الگوریتم جدید در زمینه مسائل دسته بندی میباشد که طی بررسی های انجام شده در هیچ یک از پژوهش های انجام شده در زمینه شناسایی حساب های جعلی از این الگوریتم استفاده نشده است. همانطور که مشاهده میشود این الگوریتم در بین دیگر الگوریتم های دسته بندی رتبه بالاتری را در هر پارامتر تشخیص، در زمینه تشخیص حساب های جعلی به خود اختصاص داده است که این نشان دهنده این میباشد که این الگوریتم تاثیر بالایی در شناسایی حساب های جعلی از حساب های قانونی دارا میباشد.

۴-۱- استخراج ویژگی

در این پژوهش استخراج ویژگی ها در دو مرحله و بر مبنای دو روش مبتنی بر پروفایل و مبتنی بر توییت صورت گرفته است. ابتدا با بررسی مقاله ها و پژوهش های گذشته در زمینه حساب های جعلی در تویتر بر مبنای مجموعه داده های که در این پژوهش مورد استفاده قرار گرفته است ویژگی های متناسب با مجموعه داده از مقاله ها استخراج شده اند. در مرحله دوم که تمرکز اصلی پژوهش بیشتر بر روی این قسمت میباشد و به

عبارتی ایده اصلی پژوهش میباشد به استخراج و خلق ویژگی های جدید پرداخته شده است که این امر با استفاده از ترکیب ویژگی های موجود در مجموعه داده و استفاده از محاسبات عملیاتی و منطقی با کمک زبان برنامه نویسی پایتون صورت گرفته است. از آنجا که پایتون در عملیات و بارگذاری فایل های بزرگ به مشکل برمیخورد و حافظه ی زیادی مصرف میکند، به خاطر بزرگ بودن مجموعه داده توپیت ها، آنها را به پایگاه داده SQL تبدیل و با دستور Select مقادیر را فراخوانی کرده و روی آنها عملیات و محاسبات صورت گرفته است. فرمت فایل های مجموعه داده مورد استفاده CSV است و برای افزایش قابلیت های پایتون برای کار با فایل های مجموعه داده با فرمت CSV از افزونه Pandas زبان پایتون استفاده شده است، این افزونه قابلیت های فراوانی برای انجام عملیات بر روی فایل های CSV به پایتون اضافه می کند.

۴-۲- ویژگی بر مبنای پروفایل

این ویژگی ها بر اساس مشخصات پروفایل و رفتار کاربر در شبکه های اجتماعی میباشد که شامل ویژگی های جمعیت شناختی ازجمله جزئیات مشخصات کاربری، تعداد دنبال کنندگان، تعداد دوستان، سن حساب، فرکانس توپیت، ساعات بیکاری و غیره میباشد. در این پژوهش علاوه بر ویژگی های مبتنی بر پروفایل دیگر پژوهش ها، شش ویژگی جدید استخراج شده که پنج ویژگی اول با استفاده از محاسبات منطقی و ویژگی آخر با استفاده از محاسبات عملیاتی به دست آمده اند که در ادامه در جدول ۴ شرح داده شده اند.

۴-۳- ویژگیهای مبتنی بر توپیت

ویژگی های مبتنی بر توپیت یا محتوا بر اساس پستهای متنی کاربران می باشد که بیشتر شامل تعداد هشتک ها، تعداد توپیت های حاوی آدرس، ریتوپیست، لینک های HTTP، عناوین تکراری و غیره میباشد. ویژگی های جدید مبتنی بر محتوا که در این پژوهش استخراج و استفاده شده به شرح زیر میباشند که در جدول ۴ آورده شده اند.

در این بخش در جدول ۴ به طور خلاصه تمام ویژگی های استفاده شده در این پژوهش نشان داده شده است که جدول ۴ مربوط به ویژگی های مبتنی بر پروفایل و توپیت میباشد که در این جدول، A و B, C و ویژگیهای مورد استفاده در این مقاله در نظر گرفته شده است و در کنار ویژگی های جدیدی که در این پژوهش استفاده شده اند نماد "*" قرار گرفته است. در جدول ۵ که فقط بر روی ویژگی های پروفایل تمرکز دارد، F, E و D به عنوان ویژگی های مبتنی بر پروفایل جدید که مربوط به این پژوهش می باشد.

جدول ۴: ویژگی های مبتنی بر توپیت و و پروفایل

شماره	فستهبندی	ویاگی	کارهای مرتبط
*F1	مبتنی بر پروفایل	اگر تعداد لایکها بیش از ۱۰۰ باشد.	C
*F2	مبتنی بر پروفایل	اگر طول نام کاربری کم تر از ۱۰ کاراکتر باشد.	C
*F3	مبتنی بر پروفایل	اگر موقعیت جغرافیایی فعال باشد، نام کاربری پیش فرض نباشد و کاربر دارای حداقل یک لیست باشد.	C
*F4	مبتنی بر پروفایل	اگر فیلد مکان پر شده باشد و نام کاربری و تصویر پروفایل به صورت پیش فرض نباشد.	C
*F5	مبتنی بر پروفایل	اگر کاربر از بنر پروفایل کاربری استفاده کرده باشد.	C
F6	مبتنی بر پروفایل	نسبت دنبال کنندگان به سن حساب	C
F7	مبتنی بر پروفایل	تعداد لایکها	C, B, A
F8	مبتنی بر پروفایل	تعداد دنبال کنندگان	C
F9	مبتنی بر پروفایل	تعداد دوستان	C, B
F10	مبتنی بر پروفایل	نسبت دوستان به دنبال کنندگان	C, B
F11	مبتنی بر توپیت	اگر کاربر از طریق مرورگر وب از توپیت استفاده کند.	C
*F12	مبتنی بر توپیت	مجموع هشتکها در توپیتها	C
*F13	مبتنی بر توپیت	مجموع ریتوپیستها در توپیتها	C
*F14	مبتنی بر توپیت	مجموع لایکها در توپیتها	C
*F15	مبتنی بر توپیت	مجموع Url ها در توپیتها	C
*F16	مبتنی بر توپیت	مجموع منشنها در توپیتها	C
*F17	مبتنی بر توپیت	نسبت توپیتهای ریتوپیست شده به کل توپیتها	C

*F18	مبتنی بر توییت	نسبت توییت‌های لایک شده به کل توییت‌ها	C
*F19	مبتنی بر توییت	نسبت توییت‌های دارای هشتگ به کل توییت‌ها	C
*F20	مبتنی بر توییت	نسبت توییت‌های دارای url به کل توییت‌ها	C
*F21	مبتنی بر توییت	نسبت توییت‌های دارای منشن به کل توییت‌ها	C
*F22	مبتنی بر پروفایل	دارای حداقل ۳۰ دنبال کننده	A
F23	مبتنی بر پروفایل	استفاده از موقعیت جغرافیایی	A
F25	مبتنی بر توییت	داشتن توییت‌های دارای هشتگ	A
F26	مبتنی بر توییت	ورود با iphone برای استفاده از تویتر	A
F27	مبتنی بر پروفایل	داشتن حداقل ۵۰ توییت	A
F28	مبتنی بر پروفایل	بودن در لیست دیگران	A
F29	مبتنی بر پروفایل	(number of friends) - (2*number followers)	A
F30	مبتنی بر توییت	دارای حداقل یک لیست دلخواه	A
F31	مبتنی بر توییت	منشن شده توسط دیگر کاربران	A
F32	مبتنی بر پروفایل	آیا کانت تایید شده است	B
F33	مبتنی بر توییت	نسبت تعداد ریتوییت برای هر توییت	B
F34	مبتنی بر توییت	نسبت تعداد هشتگ برای هر توییت	B
F35	مبتنی بر توییت	نسبت تعداد لایک برای هر توییت	B
F36	مبتنی بر توییت	نسبت تعداد url برای هر توییت	B
F37	مبتنی بر توییت	نسبت تعداد منشن برای هر توییت	B

جدول ۵: ویژگی های مبتنی بر پروفایل

شماره	دسته بندی	ویژگی	کارهای مرتبط
*F1	مبتنی بر پروفایل	اگر تعداد لایک‌ها بیش از ۱۰۰ باشد.	F
*F2	مبتنی بر پروفایل	اگر طول نام کاربری کم‌تر از ۱۰ کاراکتر باشد.	F
*F3	مبتنی بر پروفایل	اگر موقعیت جغرافیایی فعال باشد، نام کاربری پیش‌فرض نباشد و کاربر دارای حداقل یک لیست باشد.	F
*F4	مبتنی بر پروفایل	اگر فیلد مکان پر شده باشد و نام کاربری و تصویر پروفایل به صورت پیش‌فرض نباشد.	F
*F5	مبتنی بر پروفایل	اگر کاربر از بنر پروفایل کاربری استفاده کرده باشد.	F
F6	مبتنی بر پروفایل	نسبت دنبال کنندگان به سن حساب	F
F7	مبتنی بر پروفایل	تعداد لایک‌ها	F, D, E
F8	مبتنی بر پروفایل	تعداد دنبال کنندگان	F, D, E
F9	مبتنی بر پروفایل	تعداد دوستان	F, D, E
F10	مبتنی بر پروفایل	نسبت دوستان به دنبال کنندگان	F, D, E
F11	مبتنی بر پروفایل	سن حساب	D
F12	مبتنی بر پروفایل	$\text{friends/followers}^2$	D
F13	مبتنی بر پروفایل	تعداد توییت‌ها	D, E
F14	مبتنی بر پروفایل	دارای بیوگرافی	D
F15	مبتنی بر پروفایل	استفاده از لینک در پروفایل	D
F16	مبتنی بر پروفایل	دارای مکان	D
F17	مبتنی بر پروفایل	تعداد لیست	D, E
F18	مبتنی بر پروفایل	نسبت دوستان به سن حساب	D
F19	مبتنی بر پروفایل	$\text{friends} \geq 2 \times \text{followers}$	D
F20	مبتنی بر پروفایل	$\text{friends/followers} \approx 100$	D
F21	مبتنی بر پروفایل	بدون بیوگرافی و مکان و تعداد دوستان حداقل ۱۰۰ نفر	D
F22	مبتنی بر پروفایل	$\text{friends/followers} \geq 50$	D
F23	مبتنی بر پروفایل	تصویر پروفایل پیش‌فرض	D
F24	مبتنی بر پروفایل	استفاده از تصویر پروفایل دوم	D
F25	مبتنی بر پروفایل	نام کاربری پیش‌فرض	D
F26	مبتنی بر پروفایل	دارای تصویر پروفایل پیش‌فرض پس از ۲ ماه	D

۴-۴- داده ها

در این پژوهش برای شناسایی و تحلیل حساب های جعلی از حساب های قانونی از مجموعه داده [6] استفاده شده است.

جدول ۶: مجموعه داده های استفاده شده

مجموعه داده تویتر		
حساب های جعلی	بر اساس اطلاعات پروفایل کاربر	۳۳۵۲
	بر اساس اطلاعات توییت کاربر	۱۹۶۰۲۸
حساب های قانونی	بر اساس اطلاعات پروفایل کاربر	۳۴۵۷
	بر اساس اطلاعات توییت کاربر	۱۰۴۸۵۷۶

۴-۵- پاکسازی داده ها

در این پژوهش فایل های مربوط به ویژگی های مبتنی بر توییت و پروفایل به صورت جداگانه میباشند، برای ارزیابی نتایج نیاز به ویژگی های مبتنی بر توییت و پروفایل به صورت همزمان در مجموعه داده میباشد که نیاز به یکپارچه سازی داده ها و قرار دادن آنها در یک فایل است. بعد از این مرحله برای ارزیابی و تحلیل ویژگی ها و دست یافتن به نتایج بهتر، نیاز است تا مجموعه داده های خوب و عاری از هرگونه خطا ایجاد نماییم و از آنجا که در این پژوهش برای دست یافتن به ویژگیهای جدید از محاسبات عملیاتی استفاده شده است، ممکن است در حین محاسبات بنا بر ماهیت داده ها خطاهایی مانند، خطای تقسیم بر صفر، وجود فیلدهای خالی و غیره رخ دهد که به این منظور بر روی داده های موجود عملیات پاکسازی به صورت دستی و با استفاده از نرم افزار به کمک اپراتورهایی مانند Missing Value صورت گرفته است.

۵- پیاده سازی

در جدول ۶ تمام پارامترهای به دست آمده از الگوریتم های دسته بند این پژوهش و پژوهش های [5,6] محاسبه شده اند. همانطور که در جدول ۶ نمایش داده شده است این پژوهش توانسته است با استفاده از الگوریتم J48 در تمام پارامترهای صحت، دقت، بازخوانی و F_measure به نتایج بالای ۹۹٪ دست یابد و نسبت به دو پژوهش گذشته توانسته مقادیر این پارامترها را بهبود دهد. نتایج به دست آمده از الگوریتم درخت تصمیم نشان می دهد این پژوهش توانسته دقت تشخیص حساب های جعلی را نسبت به پژوهش های پیشین، حدود یک درصد افزایش دهد . همچنین با استفاده از الگوریتم جنگل تصادفی روش پیشنهادی توانسته در پارامترهای ارزیابی صحت، بازخوانی و F_measure به نتایج بالای ۹۸٪ و در پارامتر دقت به نتیجه ۹۹/۰۸٪ دست یابد که در مقایسه با پژوهش های [5,6] این پژوهش توانسته با الگوریتم جنگل تصادفی نتایج را ارتقا دهد. همانطور که قبلا بیان شده است الگوریتم درختهای ارتقا یافته گرادیان یک نوع الگوریتم دسته بندی جدید می باشد که فقط در این پژوهش از این الگوریتم استفاده شده است که این پژوهش توانسته با ویژگی های جدید و استفاده از این الگوریتم در مقایسه با ویژگی های استفاده شده در دو پژوهش دیگر به نتایج قابل قبولی دست یابد که طبق جدول ۵ نتایج حاکی از آن است که در تمام پارامترهای ارزیابی بهبود حاصل شده است. آخرین الگوریتم استفاده شده در این پژوهش K نزدیکترین همسایه میباشد که با اجرای این الگوریتم بر روی مجموعه ویژگی های این پژوهش در مقایسه با کارهای پیشین هیچ بهبودی در هیچکدام از نتایج حاصل نشده است. به طور کلی میتوان گفت این پژوهش در تمام پارامترها با استفاده از اکثر الگوریتم ها در مقایسه با پژوهش های پیشین بهبود داشته است و در بین الگوریتم های استفاده شده با الگوریتم جدید درخت های گرادیان ارتقا یافته که در این پژوهش از آن استفاده شده است بالاترین نتایج را به دست آورده است و با الگوریتم K نزدیکترین همسایه به ضعیف ترین نتایج دست یافته است. نتایج نشان داده شده تایید میکند که مجموعه ویژگی های جدید این پژوهش که مبتنی بر توییت و پروفایل می باشد، به طور نسبتا موثرتری حساب های جعلی در توییت را تشخیص می دهد.

جدول ۷: مقایسه بین روش ها

K-NN	Gradient Boosted Tree	Random Forest	Decision Tree	J48		الگوریتم
						پژوهش
94.46%	99.69%	98.72%	99.19%	99.50%	Accuracy	روش پیشنهادی
97.24%	99.67%	99.08%	99.34%	99.60%	Precision	
93.17%	99.80%	98.74%	99.27%	99.54%	Recall	
95.16%	99.70%	98.51%	99.17%	99.34%	F_measure	
0.993	0.999	0.999	0.995	0.997	AUC	

K-NN	Gradient Boosted Tree	Random Forest	Decision Tree	J48		الگوریتم
						پژوهش
98.30%	98.57%	98.10%	98.06%	98.68%	Accuracy	ElAzab [۵]
99.33%	98.81%	98.69%	99.07%	99.08%	Precision	
97.74%	98.74%	98.08%	97.61%	98.67%	Recall	
98.73%	98.67%	97.37%	98.26%	98.80%	F_measure	
0.999	0.997	0.996	0.995	0.992	AUC	

الگوریتم	پژوهش	J48	Decision Tree	Random Forest	Gradient Boosted Tree	K-NN
Wetstone و همکاران [۶]	Accuracy	98.68%	98.18%	98.26%	98.92%	97.64%
	Precision	99.02%	98.35%	98.81%	99.53%	97.95%
	Recall	98.74%	98.54%	98.21%	98.61%	98.01%
	F_measure	98.60%	98.25%	98.02%	99.07%	97.85%
	AUC	0.990	0.993	0.998	0.999	0.998

شکل ۴ به میزان مقادیر صحت روش پیشنهادی با پژوهش های [6,7] می باشد، همانطور که در این نمودار قابل مشاهده است روش پیشنهادی در این پژوهش توانسته است با استفاده از الگوریتم درختهای ارتقا یافته گرادیان به بیشترین مقدار صحت دست یابد که برابر با ۹۹/۳۳٪ میباشد. همچنین مشاهده میشود که این پژوهش در مقایسه با پژوهش های [6,7] با استفاده از تمام الگوریتم به جز الگوریتم جنگل تصادفی به بیشترین مقدار در دقت دست یافته است و توانسته مقدار پارامتر دقت را نسبت به پژوهش های پیشین بهبود دهد.



شکل ۴: مقایسه بین روش های [6,7] با روش پیشنهادی (سبز= روش پیشنهادی و زرد= روش ۶ و نارنجی = روش ۷)

الگوریتم J48 یکی دیگر از الگوریتم های دسته بندی مورد استفاده برای ارزیابی مجموعه داده مبتنی بر پروفایل میباشد که در شکل ۵ منحنی ROC حاصل از این الگوریتم رسم شده است. همانطور که نتایج نشان میدهد این الگوریتم توانسته به مقدار AUC برابر با ۰/۹۹۵ در روش پیشنهادی دست یابد که در مقایسه با پژوهش [7] این مقدار یکسان میباشد اما در برابر پژوهش [6] بهبود یافته است.



شکل ۵: نمودار ROC J48

۶- نتیجه گیری و جمع بندی

شبکه های اجتماعی مجازی، نسل جدیدی از فضای روابط اجتماعی اند که با اینکه عمر خیلی زیادی ندارند، به خوبی توانسته اند جای خود را در زندگی مردم باز کنند. افراد بسیاری در سنین و گروه های اجتماعی متفاوت، در شبکه های اجتماعی در کنار هم گرد می آیند و از این طریق با وجود فاصله های بسیار دور از هم در دنیای واقعی، با هم ارتباط برقرار میکنند. این شبکه ها نقش پررنگی در دنیای امروز دارند که نمیتوان آنها را نادیده گرفت، این شبکه ها بر ابعاد مختلف زندگی فردی و اجتماعی تاثیر گذارند و به همین دلیل در حال گسترش هستند و در آینده به مراتب نقش بیشتر و مهمتری را ایفا خواهند کرد. امروزه شبکه های اجتماعی به بخش جدایی ناپذیر زندگی بسیاری از مردم تبدیل شده اند. این شبکه ها با اهداف مختلفی ایجاد شده و کاربران بسیار زیادی را به خود جذب کرده اند. بعضی از شبکه های اجتماعی با سرویسهایی که ارائه میدهند و با کاربرهایی که دارند، خود به یک رسانه خبری تبدیل شده و اثرات عمیقی بر وقایع مختلف در دنیای واقعی گذاشته اند و روند زندگی افراد را تحت تاثیر قرار داده اند. هدف شبکه های اجتماعی، کمک به بهبود و تسهیل روابط، همکاری و تعامل بین افراد، سازمان ها و موسسات در موضوعات گوناگون از قبیل تجارت، پزشکی، مباحث آموزشی و غیره در راستای تغییر مثبت است. هرچند شبکه های اجتماعی مزایای فراوانی دارند اما آسیب های بسیاری نیز آنها را تهدید میکند، برخی از تهدیداتی که امنیت این شبکه ها و کاربران آنها را تحت تاثیر قرار می دهد و عرصه را برای افراد سودجو و مخرب فراهم کرده است عبارت اند از سرقت هویت، حملات فیشینگ، از دست دادن حریم خصوصی، بدافزارها، حساب های جعلی و غیره. با توجه به موارد گفته شده و اهمیت و تاثیر شبکه های اجتماعی در روند زندگی افراد در سرتاسر جهان، این پژوهش بر روی شناسایی حساب های جعلی بر روی شبکه اجتماعی محبوب و پرطرفدار توییتر تمرکز دارد. به طور کلی پژوهشهای مختلفی در این زمینه صورت گرفته است که این پژوهش ها بر مبنای سه نگرش مبتنی بر رفتار، مبتنی بر گراف و مبتنی بر دسته بندی و یادگیری ماشین انجام شده اند. در این پژوهش با انتخاب ویژگی های محاسباتی و محتوایی موثر و استفاده از رویکرد نظارتی در یادگیری ماشین و استفاده از الگوریتم های دسته بند درخت تصمیم، جنگل تصادفی، درخت های ارتقا یافته گرادیان، K نزدیکترین همسایه و الگوریتم J48 به شناسایی حساب های جعلی در توییتر پرداخته شده است. ویژگیهایی که در این پژوهش انتخاب شده اند بر مبنای ویژگی های مبتنی بر توییت و پروفایل می باشند که بر اساس مجموعه داده مورد استفاده در پژوهش تعدادی از آنها از پژوهش های پیشین استخراج شده اند و مابقی آنها شامل ویژگی های جدیدی هستند که بر اساس محاسبات عملیاتی و منطقی در این پژوهش استخراج شده اند. پس از استخراج ویژگی ها، تحلیل و ارزیابی پژوهش با پژوهش های پیشین در دو مرحله با استفاده از نرم افزار RapidMiner صورت گرفت که در ابتدا نتایج مجموعه داده هایی که شامل ویژگی های مبتنی بر پروفایل و توییت بود بررسی و ارزیابی شد و بعد از آن مجموعه داده هایی که فقط شامل ویژگیهای مبتنی بر پروفایل بود با الگوریتم های دسته بند مورد ارزیابی قرار گرفت.

۱-۶- کارهای آینده

این پژوهش توانسته با استفاده از الگوریتم درخت های ارتقا یافته گرادیان که در این پژوهش از این الگوریتم استفاده شده و در کارهای پیشین از آن استفاده نشده است بهترین نتایج را بر روی مجموعه داده توییت و پروفایل با ویژگیهای جدید به دست آورد که در بیشتر پارامترها در مقایسه با پژوهش های پیشین بهبود داشته است. همچنین در مجموعه داده مبتنی بر پروفایل با الگوریتم J48 بهترین نتایج را با استفاده از ویژگی های به دست آورده است و در اکثر پارامترها نسبت به پژوهش های پیشین بهبود یافته است.

۷- مراجع

- [1] G. Stringhini, G. Wang, M. Egele, C. Kruegel, G. Vigna, H. Zheng, et al., "Follow the green: growth and dynamics in twitter follower markets", in Proceedings of the 2013 conference on Internet measurement conference, pp. 163-176, 2013.
- [2] E. Bennett, "Social media as a tool for improving research and teaching", *Frontiers in Ecology and the Environment*, vol. 12, pp. 259-259, 2014.
- [3] K. Thomas, D. McCoy, C. Grier, A. Kolcz, and V. Paxson, "Trafficking Fraudulent Accounts: The Role of the Underground Market in Twitter Spam and Abuse", in *USENIX Security Symposium*, pp. 195-210, 2013.
- [4] M. Jiang, P. Cui, A. Beutel, C. Faloutsos, and S. Yang, "Catching synchronized behaviors in large networks: A graph mining approach", *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 10, p. 35, 2016.
- [5] E. M. Clark, J. R. Williams, C. A. Jones, R. A. Galbraith, C. M. Danforth, and P. S. Dodds, "Sifting robotic from organic text: a natural language approach for detecting automation on Twitter", *Journal of Computational Science*, vol. 16, pp. 1-7, 2016.
- [6] A. ElAzab, "Fake Accounts Detection in Twitter based on Minimum Weighted Feature", *World*, 2016.
- [7] J. H. Wetstone and S. R. Nayyar, "I Spot a Bot: Building a binary classifier to detect bots on Twitter", 2017.
- [8] A. Khalil, H. Hajjdiab, and N. Al-Qirim, "Detecting Fake Followers in Twitter: A Machine Learning Approach", 2017.