Sciences and Techniques of Information Management
Volume 8, Issue 3, 2022
Research Article

Providing a Hybrid Approach Based on Deep learning and Machine Learning to Detect Fake News - A Case Study of Persian News in the Field of COVID-19[1]

Vahid Mottaghi
PhD. Candidate in IT Management, Department of IT Management, Qeshm Branch, Islamic Azad University, Qeshm, Iran. Mvahid500@gmail.com

Mahdi Esmaeili
Assistant Professor, Department of Computer Science, Kashan Branch, Islamic Azad University, Kashan, Iran. m.esmaeili@iaukashan.ac.ir

Ghasem Ali Bazaee
Assistant Professor, Department of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran **(Corresponding author)**. Gh.bazaei@iauctb.ac.ir

Mohammad Ali Afshar Kazemi
Associate Professor, Department of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran. M_afsharkazemi@iauec.ac.ir

Abstract

Objectives: False or unconfirmed information is published on the web like accurate information, so it can become viral and influence public opinion and decisions. Fake news and gossip show the most popular forms of false and unverified information, respectively, and they should be detected as soon as possible to avoid significant effects Interest in effective identification techniques has been increasing in recent years.The problem of detecting fake news is known as a classification problem in natural language processing and text mining, and its purpose is to distinguish fake news from real and extracted texts, and to improve the accuracy of detecting fake news is the main issue of this research. Convolutional neural networks, as one of the most important models of deep learning, have gained high accuracy on these issues. These networks include problems such as not considering the position of words, which is solved by using the capsule network, and in order to achieve optimal accuracy, two problems of heavy processing of all connected layers and reducing the parametric space using the algorithm XGBOOST and particle swarm optimization (PSO) algorithm are proposed.

Methods: This study is an applied research in which about 42,000 Persian news from different cities of Iran were collected from Twitter and using additional methods of cleaning and preprocessing, additional information was removed and after tagging, the news was ready to be used for the proposed approach using Python software and related libraries are equipped with machine learning and deep learning algorithms.

1. **Received:** 2021-09-02 ; **Revised:** 2021-10-17 ; **Accepted:** 2021-12-13 ; **Published online:** 2022-09-11
DOI: 10.22091/stim.2023.2372

Published by: University of Qom.

Results: During testing, some machine learning algorithms had more power in classification problems, but with the changes in the structure of the convolutional network and Capsul network algorithm, better results were obtained than machine learning algorithms and other similar algorithms.

Conclusions: The proposed solutions in this research in comparison with the approaches of basic algorithms or solutions to solve the mentioned problems by replacing the optimal classifier and reducing the parametric space, by changing the input has been able to achieve better and more acceptable results than other approaches. And achieve an accuracy of about 96%.

Keywords: Natural Language Processing, Text Classification, Capsule Neural Networks, Fake News Detection, Corona Virus Fake News.

علوم و فنون مدیریت اطلاعات
دوره ۸، شماره ۳، ۱۴۰۱
مقاله پژوهشی

ارائه رویکرد ترکیبی مبتنی بر یادگیری عمیق و یادگیری ماشین جهت تشخیص اخبار جعلی: مطالعه موردی اخبار فارسی در حوزه کرونا ویروس[1]

وحید متقی
دانشجوی دکتری، گروه مدیریت فناوری اطلاعات، واحد قشم، دانشگاه آزاد اسلامی، قشم، ایران.
mvahid500@gmail.com

مهدی اسماعیلی
استادیار، گروه علوم کامپیوتر، واحد کاشان، دانشگاه آزاد اسلامی، کاشان، ایران (**نویسنده مسئول**).
m.esmaeili@iaukashan.ac.ir

قاسمعلی بازایی
استادیار، گروه مدیریت، واحد تهران مرکز، دانشگاه آزاد اسلامی، تهران، ایران. Gh.bazaei@iauctb.ac.ir

محمدعلی افشار کاظمی
دانشیار، گروه مدیریت، واحد تهران مرکز، دانشگاه آزاد اسلامی، تهران، ایران. M_afsharkazemi@iauec.ac.ir

چکیده

هدف: اطلاعات غلط یا تأیید نشده، دقیقاً مانند اطلاعات دقیق در وب منتشر می‌شوند. بنابراین، ممکن است ویروسی شوند و بر افکار عمومی و تصمیمات آن تأثیر بگذارند. اخبار جعلی و شایعات به ترتیب محبوب‌ترین اشکال اطلاعات دروغ و تأیید نشده را نشان می‌دهند و برای جلوگیری از تأثیرات چشمگیر آنها باید در اسرع وقت کشف شوند. علاقه به تکنیک‌های مؤثر در شناسایی، در سال‌های اخیر بسیار سریع در حال افزایش است. مسئله تشخیص اخبار جعلی به عنوان یک مسئله طبقه‌بندی در پردازش زبان طبیعی و متن‌کاوی شناخته می‌شود و هدف آن تفکیک و تشخیص اخبار جعل از واقعی، در متن‌های استخراج شده و بهبود در دقت تشخیص اخبار جعلی است. شبکه‌های عصبی کانولوشن به عنوان یکی از مهم‌ترین مدل‌های یادگیری عمیق دقت بالایی را بر روی این مسائل بدست آورده‌اند.

---

۱. **پژوهش حاضر برگرفته از:** رساله دکتری، رشته مدیریت فناوری اطلاعات، گرایش کسب‌وکار هوشمند، دانشجو: وحید متقی، با عنوان : **بهبود رویکردهای یادگیری عمیق برای مسأله تشخیص اخبار جعلی**: **مطالعه موردی اخبار فارسی درحوزه کروناویروس،** استاد راهنما: مهدی اسماعیلی و قاسمعلی بازایی، استاد مشاور: محمدعلی افشار کاظمی، ارائه شده در دانشگاه آزاد اسلامی واحد قشم است.

استناد به این مقاله: متقی، وحید؛ اسماعیلی، مهدی؛ بازایی، قاسمعلی (۱۴۰۱). ارائه رویکرد ترکیبی مبتنی بر یادگیری عمیق و یادگیری ماشین جهت تشخیص اخبار جعلی: مطالعه موردی اخبار فارسی در حوزه کرونا ویروس. *علوم و فنون مدیریت اطلاعات*، ۸(۳)، ص۲۸۳-۳۱۶.

DOI: 10.22091/stim.2021.7311.1640

تاریخ دریافت: ۱۴۰۰/۰۶/۱۱ ؛ **تاریخ اصلاح:** ۱۴۰۰/۰۷/۲۵ ؛ **تاریخ پذیرش:** ۱۴۰۰/۰۹/۲۲ ؛ **تاریخ انتشار:** ۱۴۰۱/۰۶/۲۰

ناشر: دانشگاه قم

این شبکه‌ها شامل مشکلاتی مثل عدم در نظر گرفتن موقعیت کلمات می‌باشند که مسأله مذکور با استفاده از شبکه کپسول برطرف گردیده و جهت حل مشکل پردازش سنگین لایه‌های تمام متصل و فضای پارامتریک الگوریتم‌های XGBOOST و بهینه‌سازی ازدحام انبوه ذرات (PSO) برای دستیابی به دقت و صحّت بهینه پیشنهاد شده است.

روش: مطالعه حاضر پژوهشی کاربردی بوده که در آن حدود ۴۲۰۰۰ اخبار فارسی از شهرهای مختلف ایران از توییتر جمع‌آوری شده و با استفاده از روش‌های پاک‌سازی و پیش‌پردازش، اطلاعات اضافی حذف و پس از برچسب زدن، اخبار آماده به‌کارگیری جهت رویکرد پیشنهادی با استفاده از نرم‌افزار پایتون و کتابخانه‌های مربوطه با الگوریتم‌های یادگیری ماشین و یادگیری عمیق شد.

یافته‌ها: طی بررسی، آزمایش و تست، برخی از الگوریتم‌های یادگیری ماشین دارای قدرت بیشتری در مسائل طبقه‌بندی بودند، ولی با تغییرات و اعمال روش‌های پیشنهادی که در ساختار الگوریتم شبکه کانولوشن و شبکه کپسول صورت گرفت، نتایج بهینه نسبت به الگوریتم‌های یادگیری ماشین و سایر الگوریتم‌های پایه و الگوریتم‌های مورد ارزیابی بدست آمد.

نتیجه‌گیری: راهکارهای پیشنهادی در این تحقیق در مقایسه با رویکردهای الگوریتم‌های پایه و یا راهکارهای صورت گرفته جهت حل مشکلات مذکور بدون اضافه کردن سربار اضافی از لحاظ تعداد ویژگی‌ها و عمق شبکه، با تغییر در ورودی توانسته است به نتایج بهتر و قابل قبول از سایر رویکردهای موجود در ادبیات دست یافته و به دقت و صحّت حدود ۹۶ درصد دست یابد.

کلیدواژه‌ها: پردازش زبان طبیعی، طبقه‌بندی متن، شبکه‌های عصبی کپسول، تشخیص اخبار جعل، کرونا ویروس، یادگیری عمیق، یادگیری ماشین، اخبار فارسی.

۱. مقدمه

اکنون، اینترنت به بخشی جدایی‌ناپذیر از سبک زندگی ما تبدیل شده است. نقش کانال‌های سنتی اطلاع‌رسانی مانند روزنامه‌ها و تلویزیون در پخش اخبار کم‌تر از گذشته برجسته شده است. مطمئناً، رشد رسانه‌های اجتماعی نقشی اساسی در این تحول داشته است. در واقع، شبکه‌های اجتماعی مانند Twitter[۱] و Facebook[۲] محبوبیت‌هایی را به ثبت رسانده‌اند. به عنوان مثال، فیس بوک گزارش داده است که ماه نوامبر ۲۰۱۷، ۲٫۰۷ میلیون کاربر فعال ماهانه دارد. به طور متوسط، ۱٫۳۷ میلیون از این کاربران روزانه فیس بوک ثبت‌نام می‌کنند. توییتر از ژانویه ۲۰۱۸، ۳۳۰ میلیون کاربر داشت. از زمان راه‌اندازی سیستم عامل‌های مربوطه، این تعداد دائماً در حال رشد است.

بسیاری از مردم از سیستم عامل‌های رسانه‌های اجتماعی نه تنها برای برقراری ارتباط با دوستان و خانواده، بلکه همچنین برای جمع‌آوری اطلاعات و اخبار از سراسر جهان استفاده می‌کنند. بنابراین، رسانه‌های اجتماعی نقشی اساسی در تحقق اخبار بازی می‌کنند. مطالعه موردی برای انگلیس در مون[۳] (۲۰۱۷) گزارش داده که افزایش استفاده از رسانه‌های اجتماعی و مهم‌تر از آن ارتباط آنها با مصرف اخبار است.

اخبار جعلی ممکن است یک اصطلاح نسبتاً جدید باشد، اما لزوماً پدیده جدیدی نیست (زوبیاگا و همکاران[۴]، ۲۰۱۸). با این حال، امروزه پیشرفت‌های فناوری و انتشار اخبار از طریق رسانه‌های مختلف باعث گسترش اخبار جعلی شده است. به همین ترتیب، نرخ اخبار جعلی اخیراً به صورت تصاعدی افزایش یافته است و باید کاری انجام شود تا از ادامه این کار در آینده جلوگیری شود. کشف اخبار جعلی وظیفه طبقه‌بندی اخبار با توجه به صحّت آن است. در یک نگاه ساده، این یک کار طبقه‌بندی باینری است، در حالی که در یک محیط متفاوت، این یک کار طبقه‌بندی چندگانه نیز به‌شمار می‌آید (زوبیاگا و همکاران، ۲۰۱۸).

یادگیری عمیق به دلیل انتخاب ویژگی‌ها به صورت خودکار در تشخیص اخبار جعلی به شدت

---

۱. توییتر

۲. فیس بوک

3. Moon

4. Zubiaga & et al.

مورد استفاده قرار می‌گیرد. شبکه‌های عصبی کانولوشن از الگوریتم‌های یادگیری عمیق بوده و انتخاب ویژگی‌ها در آن به صورت خودکار می‌باشد. این رویکرد دارای مشکلاتی نظیر عدم در نظر گرفتن موقعیت کلمات در جمله، وجود لایه‌های تماماً متصل و فضای پارامتریک می‌باشند. هدف کلی پژوهش حاضر ارائه روش‌هایی برای بهبود این مشکلات است که برای این منظور رویکردهای متفاوتی معرفی می‌شوند.

در تحقیق پیش‌رو برای حل مشکل عدم در نظر گرفتن موقعیت کلمات، ترکیب شبکه کپسول با Bi-GRU پیشنهاد می‌شود و همچنین با جای‌گذاری یک طبقه‌بندی XGBoost سعی در حذف لایه‌های تماماً متصل داریم. الگوریتم PSO[1] برای بدست آوردن فضای پارامتریک بهینه نیز در ادامه کار مورد استفاده قرار می‌گیرد.

۲. چارچوب نظری و پیشینه پژوهش

به اعتقاد پژوهشگران (آلکات و گنزو[2]، ۲۰۱۷)، رسانه‌های اجتماعی به ابزاری مهم برای انتشار روزنامه‌نگاران تبدیل شده است و روش اصلی مصرف برای شهروندانی است که به دنبال آخرین اخبار هستند. روزنامه‌نگاران ممکن است از رسانه‌های اجتماعی برای گزارش در مورد افکار عمومی در مورد انتشار اخبار و حتی برای کشف داستان‌های جدید بالقوه استفاده کنند، در حالی که شهروندان ممکن است توسعه اخبار و رویدادهای جدید را از طریق کانال‌های رسمی (یعنی حساب‌های رسمی رسانه‌های خبری در سیستم عامل‌های رسانه‌های اجتماعی) یا از طریق پست‌های شبکه خودشان (به عنوان مثال دوستان، خانواده، شخصیت‌های عمومی) دریافت کنند. در واقع، شبکه‌های اجتماعی به ویژه در شرایط بحرانی، به دلیل توانایی ذاتی آنها در انتشار اخبار فوری، بسیار سریع‌تر از رسانه‌های سنتّی، عمل می‌کنند.

وقایع سیاسی اخیر منجر به محبوبیت و گسترش اخبار جعلی شده است. برخی از این اخبارها به گونه‌ای گسترش پیدا می‌کنند که به واقعیت بسیار نزدیک هستند و بسیاری از افراد را تحت تأثیر قرار می‌دهند. روش‌های مختلفی برای خودکارسازی روند کشف اخبار جعلی ارائه شده‌اند، که از محبوب‌ترین اینگونه تلاش‌ها می‌توان به لیست‌های سیاه از منابع و نویسندگان غیرقابل اعتماد

1. Particle Swarm Optimization
2. Allcott & Gentzkow

اشاره کرد. در حالی که این ابزارها مفید هستند، برای ایجاد یک راه‌حل کامل‌تر برای کشف چنین اخبارهایی، باید موارد بیشتری در نظر گرفته شود. روش‌های یادگیری ماشین و تکنیک‌های پردازش زبان طبیعی، ابزارهای مفیدی برای تشخیص اخبار جعل هستند (ژانگ و همکاران[1]، ۲۰۱۲؛ رابین و همکاران[2]، ۲۰۱۶).

گرچه مشکل اخبار جعلی موضوع جدیدی نیست و می‌توان استدلال کرد که از ابتدای گسترش روزنامه‌های کاغذی وجود داشته است، انسان‌ها همواره تمایل دارند که اطلاعات غلط را باور کنند. اخبار جعلی در چند سال گذشته، به ویژه پس از انتخابات ایالات متحده در سال ۲۰۱۶، مورد توجه بیشتری قرار گرفته است. تشخیص اخبار جعلی برای انسان‌ها دشوار است. می‌توان ادعا کرد که تنها راه شناسایی آنها در صورت عدم داشتن داده‌های زیاد، به صورت دستی است (باندیلی و مارسلونی[3]، ۲۰۱۹).

اکنون اخبار جعلی به عنوان یکی از بزرگ‌ترین تهدیدات برای دموکراسی، روزنامه‌نگاری و آزادی بیان تلقی می‌شود و اعتماد عمومی را به دولت‌ها تضعیف می‌کند. تأثیر احتمالی آن در همه‌پرسی بحث برانگیز "Brexit" و انتخابات ریاست جمهوری آمریکا در سال ۲۰۱۶ مورد اهمیت قرار گرفت (باندیلی و مارسلونی، ۲۰۱۹). اقتصاد ما نیز از انتشار خبرهای جعلی مصون نیست، زیرا اخبار جعلی به نوسانات بورس سهام و معاملات گسترده مرتبط است. به عنوان مثال، اخبار جعلی که ادعا می‌کرد باراک اوباما در اثر انفجار زخمی شده است، ۱۳۰ میلیارد دلار ارزش سهام را از بین برد (باندیلی و مارسلونی، ۲۰۱۹). این وقایع و ضررها باعث ایجاد انگیزه در تحقیقات خبری جعلی شده و بحث پیرامون اخبار جعلی را برانگیخته است.

بدیهی است که رایج‌ترین اصطلاحات در مراجع اصلی اخبار جعلی و شایعات هستند، اما با این وجود محققان جنبه‌های دیگری را که مربوط به اطلاعات غلط در وب هستند، از قبیل clickbait، اسپم‌های اجتماعی و بررسی‌های جعلی نیز مورد تجزیه و تحلیل قرار داده‌اند. در ادبیات، طبقه‌بندی‌های مختلفی از اخبار جعلی و شایعات ارائه شده است، بیشتر به منبع و نوع داده‌های مورد استفاده برای تجزیه و تحلیل بستگی دارد. مطالعات اولیه در این زمینه، به ویژه از

1. Zhang & et al.
2. Rubin & et al.
3. Bondielli & Marcelloni

منظر محاسباتی، در چند سال اخیر انجام شده است. بنابراین، مرزهای مطالعه اخبار جعلی اغلب به روشنی مشخص نشده است. به همین دلیل، ما معتقدیم که ارائه بینش در مورد اینکه چه نوع داده‌ای می‌تواند تبدیل به موضوع تجزیه و تحلیل و چگونگی تعریف آن شود، اساسی است. در ادامه دو نوع از این اطلاعات آورده شده است.

● **اخبار جعلی:** واژه اخبار جعلی برای شناسایی اطلاعات نادرست در رسانه‌های اصلی، به ویژه برای مطالب مرتبط با وب، مطرح شده است. با این حال، تحقیق در مورد اخبار جعلی به طور کلی از تعریف محدودتری استفاده می‌کند. در پژوهشی باندیلی و مارسلونی (۲۰۱۹) یک خبر جعلی را در یک مقاله خبری عمداً و با صحت اثبات دروغ به‌کار بردند. چنین تعریفی به دو جنبه اصلی قصد و صحت وابسته است. بنابراین، اخبار جعلی مقالاتی خبری هستند که عمداً برای گمراه کردن یا سوءاستفاده از خوانندگان نوشته شده‌اند، اما می‌توان آنها را با استفاده از منابع دیگر تأیید کرد. چندین مطالعه اخیر، مانند کانروی و همکاران[1] (۲۰۱۵)، این تعریف را تصویب کرده‌اند. در پژوهش رابین و همکاران[2] (۲۰۱۵) تمایز میان جنبه‌های مختلف اخبار جعلی معرفی شده است. به طور خاص، نویسندگان بر جعل جدی، جعل در مقیاس بزرگ و تقلب‌های طنز توجه کرده‌اند. جعل‌های جدی نمونه اولیه اخبار جعلی می‌باشند، یعنی مقالاتی با قصد مخرب هستند (مثلاً مصاحبه‌های جعلی، مقالات شبه علمی و غیره) که اغلب از طریق رسانه‌های اجتماعی ویروسی می‌شوند. جعل در مقیاس بزرگ، گزارشی از اطلاعات دروغین است که به عنوان خبرهای مناسب پنهان شده‌اند (رابین و همکاران، ۲۰۱۵).

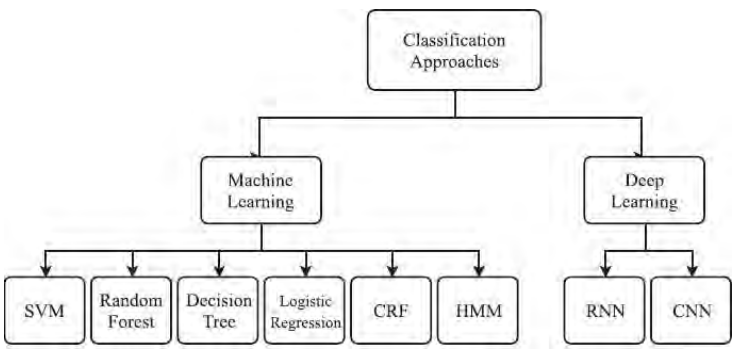
● **شایعات:** در ادبیات علمی اخیر، شایعات احتمالاً بیشترین اطلاعات دروغ مورد مطالعه در وب بوده‌اند. آنها به اطلاعاتی اشاره می‌کنند که هنوز توسط منابع رسمی تأیید نشده‌اند و بیشتر توسط کاربران در سیستم عامل‌های رسانه‌های اجتماعی پخش شده‌اند.

شایعات، محصول عصر اینترنت نیستند، با مطالعات اولیه مربوط به پایان جنگ جهانی دوم می‌توان دریافت که همواره شایعات وجود داشته‌اند (آلپورت و پستمن[3]، ۱۹۴۶؛ میر[4]، ۱۹۶۹). با

1. Conroy & et al.
2. Rubin & et al.
3. Allport & Postman
4. Meyer

این حال، می‌توان استدلال کرد که اینترنت و بسترهای رسانه‌های اجتماعی به طور خاص زمینه باروری برای انتشار اطلاعات غیرقابل اثبات و شایعات هستند (وثوقی و همکاران[1]، ۲۰۱۷).

با توجه به تعداد وظایف فرعی مختلف ارائه شده در ادبیات، انجام یک ارزیابی مطمئن و منصفانه از رویکردهای کشف اطلاعات غلط و مقایسه عملکرد آنها بسیار سخت است. علاوه بر این، غالباً محققان برای یافتن مناسب‌ترین مدل برای مجموعه داده‌ها و کارها، مجبور به آزمایش الگوریتم‌های طبقه‌بندی مختلف هستند. شکل (۱) رویکردهای مختلفی را که ما در تحلیل خود در نظر خواهیم گرفت، نشان می‌دهد. این رویکردها را می‌توان به رویکردهای طبقه‌بندی و رویکردهای دیگر دسته‌بندی کرد. رویکردهای طبقه‌بندی می‌توانند به نوبه خود مبتنی بر یادگیری ماشین و یادگیری عمیق باشند.



شکل ٦- رویکردهای موجود در ادبیات برای مسأله اخبار جعل (باندیلی و مارسلونی، ۲۰۱۹)

ثابت شده است که الگوریتم‌های یادگیری ماشین برای حل بسیاری از کارها در زمینه مهندسی اطلاعات می‌توانند مفید واقع شوند. از آنجایی که اولین رویکردها با تمرکز بر اعتبار در رسانه‌های اجتماعی (کاستیلو و همکاران[2]، ۲۰۱۱) و کشف فریب با واسطه رایانه (کاستیلو و همکاران، ۲۰۱۱)، نتایج امیدوارکننده‌ای را ارائه داده‌اند، تکنیک‌های یادگیری ماشین در تعدادی از تحقیقات در رابطه با مسئله کشف اطلاعات غلط مورد تست و ارزیابی قرار گرفتند. بیشتر رویکردهای یادگیری ماشین که برای اخبار جعلی و کشف شایعات به کار رفته‌اند، از یک استراتژی یادگیری

1. Vosoughi & et al.
2. Castillo & et al.

نظارت‌شده استفاده می‌کنند. در جدول شماره (۱) به برخی از این الگوریتم‌ها و تلاش‌ها برای کشف اخبار جعلی اشاره شده است.

جدول ۵- دسته‌بندی رویکردهای یادگیری ماشین سنّتی برای اخبار جعل

روش	منبع
Support Vector Machine (SVM)	(ژانگ و همکاران، ۲۰۱۲؛ رابین و همکاران، ۲۰۱۶؛ کاستیلو و همکاران، ۲۰۱۱؛ کین و همکاران[۱]، ۲۰۱۶؛ یانگ و همکاران[۲]، ۲۰۱۲)
Decision Tree (DT)	(ژو و همکاران[۳]، ۲۰۰۳؛ چانگ و همکاران[۴]، ۲۰۱۶؛ کاستیلو و همکاران، ۲۰۱۱؛ آکر و همکاران[۵]، ۲۰۱۷؛ روچانسکی و همکاران[۶]، ۲۰۱۷)
Random Forest (FR)	(ژو و همکاران، ۲۰۰۳؛ آکر و همکاران، ۲۰۱۷؛ بریسکو و همکاران[۷]، ۲۰۱۴؛ زوبیاگا و همکاران، ۲۰۱۶؛ کوون و همکاران[۸]، ۲۰۱۳)
Logistic Regression (LR)	(ژو و همکاران، ۲۰۰۳؛ جایسمیدیس و همکاران[۹]، ۲۰۱۶؛ زنگ و همکاران[۱۰]، ۲۰۱۶)
Conditional Random Field (CRF)	(زوبیاگا و همکاران، ۲۰۱۶)
Hidden Markov Models (HMMs)	(وثوقی و همکاران، ۲۰۱۷؛ وثوقی، ۲۰۱۵)

۲-۱. رویکردهای مبتنی بر یادگیری عمیق

یادگیری عمیق[۱۱] یکی از مباحث تحقیقاتی است که به طور گسترده در زمینه یادگیری ماشین مورد بررسی قرار می‌گیرد. طبقه‌بندهای Deep Learning[۱۲] به دلیل نتایج بسیار امیدوارکننده در تعدادی از زمینه‌های تحقیقاتی از جمله متن کاوی و NLP[۱۳]، در سال‌های اخیر به شدت مورد

1. Qin & et al.
2. Yang & et al.
3. Zhou & et al.
4. Chang & et al.
5. Aker & et al.
6. Ruchansky & et al.
7. Briscoe & et al.
8. Kwon & et al.
9. Giasemidis & et al.
10. Zeng & et al.
11. Deep Learning

۱۲. یادگیری عمیق

13. Neuro Linguistic Programming

استفاده قرار گرفته‌اند. چارچوب‌های یادگیری عمیق به دلیل انتخاب ویژگی‌ها به صورت خودکار، دقت بالاتری را نسبت به رویکردهای یادگیری سنتی بدست آورده‌اند. وظیفه استخراج ویژگی وقت‌گیر بوده و ممکن است منجر به ویژگی‌های مغرضانه شود (ما و همکاران[1]، ۲۰۱۶). این یک موضوع مهم برای کارهایی مانند اخبار جعلی و کشف شایعات است، جایی که شناسایی ویژگی‌های مربوط به آنالیز ممکن است یک چالش بزرگ‌تر را ایجاد کند. از سوی دیگر، چارچوب‌های یادگیری عمیق می‌توانند بازنمایی‌های پنهان را از ورودی‌های ساده‌تر بدست بیاورند (ما و همکاران، ۲۰۱۶). بنابراین، مسئله از مدل‌سازی ویژگی‌های ورودی مرتبط، به مدل‌سازی خود شبکه منتقل می‌شود که امکان حل کارآمد مسئله را فراهم می‌آورد. دو الگوریتم که در شبکه‌های عصبی مصنوعی مدرن بسیار مورد استفاده قرار می‌گیرند، شبکه‌های عصبی مکرر (RNN)[2] و شبکه‌های عصبی کانولوشنی (CNN)[3] هستند.

جدول شماره (۲) رویکردهای موجود در ادبیات را که از شبکه‌های عصبی عمیق استفاده کرده‌اند، نشان می‌دهد:

جدول ۲- رویکردهای شبکه عصبی عمیق

روش	منبع
Convolutional Neural Network (CNN)	(جاکووی و همکاران[4]، ۲۰۱۸؛ لی کان و همکاران[5]، ۲۰۱۰؛ گارل و همکاران[6]، ۲۰۱۸؛ چن و همکاران[7]، ۲۰۱۷)
Recurrent Neural Networks (RNNs)	زوبیاگا و همکاران، ۲۰۱۶؛ بریسکو و همکاران، ۲۰۱۴؛ کوچکینا و همکاران[8]، ۲۰۱۸؛ گارل و همکاران، ۲۰۱۸؛ ما و همکاران، ۲۰۱۶؛ روچانسکی و همکاران، ۲۰۱۷؛ چو و همکاران[9]، ۲۰۱۴)

1. Ma & et al.
2. Recurrent Neural Network
3. Convolutional neural networks
4. Jacovi & et al.
5. Lecun & et al.
6. Gorrell & et al.
7. Chen & et al.
8. Kochkina & et al.
9. Cho & et al.

۳. سؤالات پژوهش

۱) آیا در این تحقیق رویکردهای یادگیری ماشین، دقت و صحت بهتری نسبت به رویکردهای یادگیری عمیق در تشخیص اخبار جعلی فارسی دارند؟

۲) آیا استفاده از شبکه‌های کپسول مشکل آنالیز در سطح جملات حل شده و طبقه‌بندی متن‌های فارسی امکان‌پذیر بوده و موجب بهبود دقت و صحت تشخیص اخبار جعلی شده است؟

۳) آیا با تلفیق شبکه کپسول با الگوریتم XGBOOST و PSO مشکل پردازش لایه‌های تماماً متصل و فضای پارامتریک حل شده و نتایج حاصل از دقت و صحت بهینه برخوردار است؟

۴. روش تحقیق

شکل ۲- گردش کاری برای هر پروژه پردازش زبان طبیعی به صورت کلی

۴-۱. مجموعه داده‌های جمع‌آوری شده اخبار فارسی

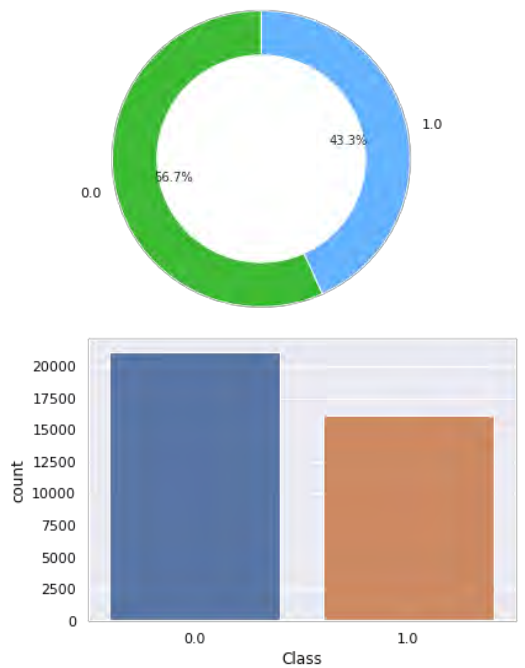
داده‌های جمع‌آوری شده در این پژوهش به مدت ۲ ماه از تاریخ ۱۳۹۹/۰۱/۰۱ لغایت ۱۳۹۹/۰۳/۰۵ به تعداد ۴۲۰۰۰ خبر فارسی بوده که از این داده‌ها ۸۰ درصد برای آموزش و ۲۰ درصد برای تست مدل استفاده شده است. این داده‌ها در چندین مرحله جمع‌آوری و تکمیل شده است. در هر مرحله اطلاعات بیشتر به داده‌های اولیه اضافه شده است. داده‌ها در سطرهای مجزا ذخیره گردیده که هر سطر را می‌توان متناظر با یک کاربر و یک داده ورودی برای آموزش مدل در نظر گرفت. در ادامه به ترتیب از جمع‌آوری داده‌های اولیه شروع می‌کنیم و سپس در هر مرحله تکمیل داده‌های اولیه و اضافه شدن ویژگی‌های مورد نیاز را بررسی می‌کنیم و در نهایت اطلاعات آماری کل داده‌های جمع‌آوری شده مطرح می‌شود.

توییتر اغلب به عنوان یک پلتفرم عمومی شناخته می‌شود و API های متفاوتی برای جمع‌آوری داده‌های آن معرفی شده است (میلر، ۲۰۱۷). در این کار از GetOldTweets و Tweepy برای جمع‌آوری داده‌ها استفاده شد. بعضی از ابزارها دسترسی به توییت‌های قدیمی‌تر را فراهم می‌کنند. GetOldTweets یک ابزار کاملاً رایگان برای جمع‌آوری داده‌های توییتر است که از خاصیت‌های جستجوی ترکیبی و جستجو بر اساس یک کلمه نیز پشتیبانی می‌کند و این اجازه را می‌دهد تا به توییت‌های طولانی مدت دست یابیم. این API اطلاعات بسیار مفیدی نظیر

id (str), permalink (str), username (str), to (str), text (str), date (datetime) in UTC, retweets (int), favorites (int), mentions (str), hashtags (str), geo (str) را فراهم می‌آورد. ویژگی‌هایی که GetOldTweets استخراج می‌کند، مفید اما بسیار کم می‌باشند. برای این منظور ما ازTweepy برای استخراج برخی از ویژگی‌های مفید دیگر نظیر تعداد followerها و following های هر فرد استفاده می‌کنیم. این ابزار قدرتمند نیز برای جمع‌آوری داده‌های توییتر مورد استفاده قرار می‌گیرد که از مکانیزم OAuth برای احراز هویت استفاده می‌کند. برای جمع‌آوری داده‌ها چهار شهر مختلف تهران، اصفهان، قم و مشهد به عنوان منطقه مکانی در نظر گرفته شد. همچنین دو هشتگ #کرونا و #کرونا_ویروس به عنوان هشتگ جمع‌آوری در نظر گرفته شد.

۴-۲. برچسب زدن

در این بخش داده‌ها توسط گروهی از افراد خبره و با کمک سایت‌های خبری برچسب مناسب زده شده است. برچسب صفر به عنوان اخبار واقعی و برچسب یک، به عنوان اخبار جعلی محسوب می‌گردد.



شکل ۳- آنالیز آماری مربوط به توییت‌های جمع‌آوری شده
(در این دو نمودار ۰ نشان‌دهنده کلاس Real و ۱ نشان‌دهنده کلاس Fakeمی‌باشد)

۴-۳. فاز پیش‌پردازش (استخراج متون و پیش‌پردازش)

در این گام پیش پردازش‌های رایج در حوزه پردازش متن نظیر حذف ایست واژه‌ها، حذف علائم نگارشی و توکن‌بندی بر روی متن‌های برچسب دار اعمال شده است. معمولاً چند مرحله در زمینه پاک‌سازی و پیش‌پردازش داده‌های متنی وجود دارد. با این حال در این بخش نیز برخی از مهم‌ترین گام‌هایی را که در تحقیق پیش‌رو صورت گرفته به اختصار توضیح می‌دهیم. این گام‌ها به وفور در پروژه‌های NLP مورد بهره‌برداری قرار می‌گیرند. در این تحقیق از nltk و scikit-learn و spacy استفاده می‌کنیم که جزء کتابخانه‌های تثبیت‌شده در حوزه NLP هستند.

۴-۳-۱. حذف تگ‌های HTML

متن‌های ساخت‌نیافته غالباً شامل مقدار زیادی نویز هستند، به خصوص اگر از تکنیک‌هایی مانند اسکراپ کردن وب یا صفحه استفاده کنید. تگ‌های HTML[1] به طور معمول یکی از مؤلفه‌هایی هستند که ارزش زیادی در جهت درک و آنالیز متن اضافه نمی‌کنند.

۴-۳-۲. کاراکترهای ویژه

کاراکترهای ویژه و نمادها معمولاً کاراکترهای عددی- حرفی یا حتی در مواردی کاراکترهای عددی (بسته به مسئله) باعث افزایش نویز در متون می‌شوند که بایستی حذف گردد.

۴-۳-۳. حذف ایست‌واژه‌ها[2]

کلماتی که بی‌معنی هستند یا معنای خاصی ندارند، به خصوص وقتی ویژگی‌های معنایی از متن استخراج می‌شود، به نام ایست‌واژه نامیده می‌شوند. این موارد معمولاً کلماتی هستند که فراوانی زیادی در متن دارند. به طور معمول این کلمات شامل حروف تعریف، ربط، اضافه و مواردی از این دست هستند.

۴-۳-۴. نرمال‌سازی متن

با جمع‌بندی داده‌ها و مرتبط ساختن عملیات‌هایی همچون موارد مذکور، یک نرمال‌سازی متن برای پیش‌پردازش داده‌های متنی ایجاد می‌کنیم.

از دستور CountVectorizer که توسط کتابخانه scikit-learn برای برداری‌سازی جملات

1. Hyper Text Markup Language
2. Stopwords

فراهم شده، استفاده می‌گردد. این پیاده‌سازی کلمات هر جمله را دریافت کرده و یک واژه از همه کلمات یکتای موجود در جمله می‌سازد. این واژه را می‌توان برای ساخت یک بردار ویژگی از شمارش‌های کلمات، مورد استفاده قرار داد.

CountVectorizer کار «توکن‌سازی»[1] را انجام می‌دهد و به عبارت دیگر، جملات را به مجموعه‌ای از «توکن‌ها» جداسازی می‌کند. علاوه بر این، نشانه‌گذاری‌ها و کاراکترهای خاص را نیز حذف کرده و می‌تواند پیش‌پردازش را بر هر واژه اعمال کند.

۴-۴. جاسازی کلمات

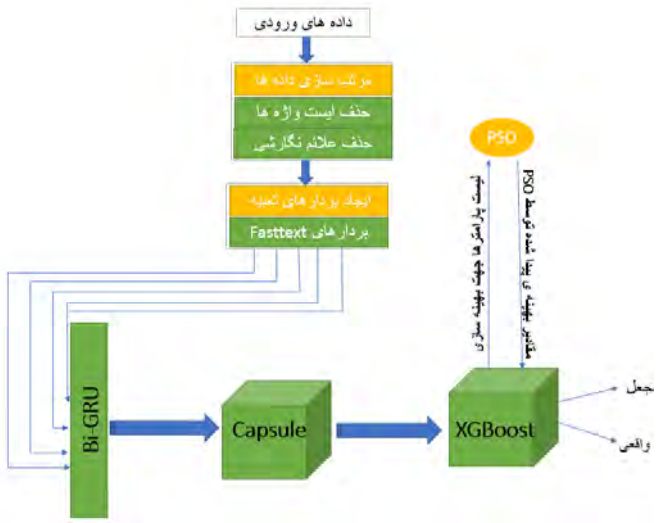
یکی از مشکلات متن‌ها برای شبکه‌های عصبی این است که نمی‌توان آن‌ها را به صورت مستقیم به شبکه داد. برای این منظور آن‌ها را به صورت بردار تعبیه شده به شبکه می‌فرستند. راه‌های مختلف بردارسازی متن به صورت کلمات و کاراکترها بوده که در این پژوهش پس از بردارسازی با روش (fasttext) که نوعی جاسازی کلمات از پیش تعیین شده بوده و توسط شرکت فیس بوک ابداع شده، اقدام گردیده که به آن فضای جاسازی[2] گفته می‌شود.

۴-۵- آموزش و تست مدل: تعریف مدل مبنا

هنگام انجام کار یادگیری ماشین، یک گام مهم تعریف مدل مبنا است. مدل مبنا معمولاً شامل یک مدل ساده بوده که بعداً برای مقایسه با مدل پیشرفته‌تری که قصد آزمودن آن وجود دارد، مورد استفاده قرار می‌گیرد. در این شرایط، از مدل مبنا برای مقایسه آن با متدهای پیشرفته‌تر از جمله شبکه‌های عصبی (عمیق) بهره‌برداری می‌شود. ابتدا، باید داده‌ها به مجموعه‌های «آموزش[3]» و «آزمون[4]» تقسیم شوند تا امکان ارزیابی صحت و قابلیت تعمیم مدل فراهم شود.

پس از توضیحات تکمیلی فوق، تمامی مراحل کار به صورت چارت نمایش داده شده (شکل چارچوب پیشنهادی) و به رویکرد پیشنهادی بیان شده، استفاده از شبکه کپسول و تلفیق الگوریتم XGBOOST و PSO که ایده اصلی تحقیق است، می‌پردازیم.

1. Tokenization
2. Embedding Space
3. Train
4. Test



شکل ٤- چارچوب روش پیشنهادی

۵. رویکرد پیشنهادی Bi-GRU Capsule

مسئله تشخیص اخبار جعلی به عنوان یک مسئله طبقه‌بندی در پردازش زبان طبیعی و متن‌کاوی شناخته می‌شود و هدف آن اختصاص یکی از کلاس‌های جعل و واقعی به متن‌های آزاد می‌باشد. شبکه‌های عصبی کانولوشن به عنوان یکی از مهم‌ترین مدل‌های یادگیری عمیق دقت بالایی را بر روی این مسائل بدست آورده است. در این شبکه‌ها کلمات به صورت کیسه‌ای از کلمات به مدل داده می‌شوند که هر کلمه با توجه به فضای برداری (Glove, Word2vec, FastText) به ماتریس‌های دو بعدی تبدیل می‌شوند. یکی از مهم‌ترین چیزهایی که در این شبکه‌ها از بین می‌رود، مکان قرار گیری کلمات در سطح جملات است.

نکته قابل توجه این است که در شبکه‌های عصبی کانولوشن، انتخاب ویژگی‌ها به صورت خودکار است. این شبکه‌ها در سطح کلمه کار می‌کنند و نمی‌توانند رابطه و فاصله بین جملات را در نظر بگیرند. برای این منظور نیاز است که این رابطه‌ها در نظر گرفته شود، چرا که ممکن است این رابطه‌ها دارای ویژگی‌های مفیدی باشند.

در نظر گرفتن کل جملات به عنوان دنباله‌ای از کلمات ورودی در شبکه‌های کانولوشن باعث می‌شود تأثیر ترتیب کلمات به خوبی در نظر گرفته نشود.

برای حل این مشکل صبور و همکاران (۲۰۱۷)، شبکه‌های کپسول را پیشنهاد دادند. در واقع یکی دیگر از محدودیت‌های CNN آن است که نرون‌های آن بر اساس احتمال ویژگی‌های خاصی

(و نه همه ویژگی‌ها) فعال می‌شوند. بنابراین، بر روی روابط این ویژگی‌ها آموزش نمی‌بینند. تعیین روابط بین ویژگی‌ها در CNN نیازمند اطلاعات بیشتری است که ممکن است این اطلاعات، حتی در صورت وجود در لایه‌های اول، در لایه‌های pooling از بین رفته باشد. به بیان دیگر، نمایش داده‌ها داخل یک شبکه، سلسله مراتب روابط بین داده‌ها را لحاظ نمی‌کند. شبکه‌های کپسول جهت حل این مشکلات ارائه شده‌اند. کپسول به معنی پوشش و محافظ است و به این دلیل استفاده شده که توسط این روش تمام ویژگی‌های مهم داده‌های ورودی حفظ می‌شود. تفاوت مهم شبکه‌های کپسول با کانولوشن در آن است که برخلاف کانولوشن که مقادیر به صورت عددی (اسکالر) ذخیره می‌شوند، در کپسول به شکل بردار ذخیره می‌گردند. هر بردار در ذات خود دارای دو ویژگی اندازه و جهت است. در کپسول، یک ویژگی خاص با اندازه بردار و تغییر آن با جهت بردار مدل می‌شود. بنابراین، با تغییر پارامترهای یک ویژگی خاص، اندازه بردار (خود ویژگی) ثابت، ولی جهت آن (پارامترها) تغییر می‌کند (صبور و همکاران، ۲۰۱۷).

در شبکه‌های کپسول، محاسبات پیچیده‌ای بر روی ویژگی‌ها انجام می‌شود و نتایج حاصل از این محاسبات به یک بردار حاوی اطلاعات بسیار مفید نگاشت می‌گردد. بر این مبنا هر کپسول می‌آموزد که موجودیت‌های (ویژگی‌های) ظاهری را در یک دامنه محدود از شرایط مشاهده و تغییر شکل دهد و دو احتمال، یکی وجود ویژگی‌ها در یک دامنه محدود و دیگری مجموعه‌ای از پارامترهای اکتشافی مرتبط با این ویژگی را به عنوان خروجی تولید کند. در صورتی که کپسول درست کار کند، احتمال وجود خود ویژگی اصلی غیرقابل تغییر invariant است، در حالی که احتمال پارامترهای اکتشافی مرتبط با این ویژگی‌ها، متغیر equivalent می‌باشد. هر چند این شبکه در پژوهش نگوین و همکاران (۲۰۱۹) برای تشخیص اخبار جعلی بر روی داده‌های تصویری و ویدئویی مورد استفاده قرار گرفته است، اما در مورد داده‌های متنی مورد تست قرار نگرفته‌اند. انتظار داریم با این شبکه‌ها بتوانیم تأثیر فاصله کلمات را به عنوان ویژگی برای بهبود رویکرد طبقه‌بندی در نظر بگیریم.

رویکرد پیشنهادی ما بسیار شبیه به پژوهش راتنایکا و همکاران (۲۰۱۸) می‌باشد، با این تفاوت که ما سعی می‌کنیم که از فراداده‌ها برای بهبود رویکرد پیشنهادی استفاده کنیم. رویکرد پیشنهادی شامل چهار لایه اساسی است که به صورت زیر می‌باشند:

۱) لایه Words embedding: در این لایه هر یک از اسناد طبق الگوی گفته شده در بخش قبل، به بردارهای چگال تبدیل می‌شوند، که این بردارهای چگال توسط Glove Words embedding به

دست می‌آیند. برای این کار از مجموعه پیش‌آموزش داده‌شده Glove[1] با ۲۰۰ بُعد استفاده شد. خروجی این لایه به ازای هر یک از اسناد برابر مقدار مقابل می‌باشد: $out_{embed} = M * E$

که در آن M برابر بیشترین طول سند در کل مجموعه داده‌ها و E اندازه بردارهای تعبیه شده است که ما از Glove با اندازه ۲۰۰ استفاده کردیم.

۲) لایه Bi-GRU: ورودی لایه Bi-GRU بردارهای تعبیه شده‌ای هستند که از لایه embedding گرفته شده‌اند. اگر این بردارها را به صورت $Out_{embed} = [X_1, X_2, \dots, X_n]$ نمایش دهیم، ورودی Bi-GRU در گام T بردار $X \in R^{200}$ می‌باشد. آنگاه $h_t = h_1, h_2, \dots, h_t$ نشان‌دهنده دنباله بردارهای مخفی[2] در Bi-GRU می‌باشد و از طریق رابطه زیر محاسبه می‌شوند:

$$z_t = \sigma(w_z x_t + U_z h_{t-1} + b_z)$$

$$r_t = \sigma(w_r x_t + U_r h_{t-1} + b_r)$$

$$h'_t = \sigma(w_h x_t + U_n(r_t \odot h_{t-1}) + b_h)$$

$$h_t = (1 - z)h_{t-1} + z_t h'_t$$

که در آن z_t گیت update ، r_t گیت rest، h'_t گیت شرطی و h_t خروجی فعال می‌باشد. $w_z, w_r, W_n, U_z, U_r, U_n$ماتریس‌های قابل یادگیری، b_n, b_z, b_r بایاس‌های قابل یادگیری، σ تابع فعال‌ساز و $\odot$ علامت ضرب نقطه‌ای بین عناصر می‌باشد. شبکه‌های Bi-GRU در حالت عادی داده‌ها را در یک جهت مدل می‌کنند. در برخی از وظایف معکوس دنباله ورودی می‌تواند عملکرد شبکه را بهبود بخشد. شبکه‌های Bidirectional GRU (Bi- GRU) (شوستر و پالیوال[3]، ۱۹۹۷) داده‌ها را در هر دو جهت Forward و Backward پردازش می‌کنند. اگر $\overrightarrow{h_t}$ خروجی Forward برای دنباله x_1^t باشد، $t = 1, 2, 3, \dots, t$ می‌باشد و اگر $\overleftarrow{h_t}$ خروجی Backward به ازای $t = t, \dots 3, 2, 1$ می‌باشد. آن‌گاه خروجی Bi-GRU از طریق ترکیب گام به گام خروجی‌های Forward و Backward به صورت $h_t = (\overrightarrow{h_t}. \overleftarrow{h_t})$ به دست می‌آید. در مقایسه با حالت unidirectional این شبکه دو برابر بیشتر پارامتر آزاد دارد.

۳) لایه کپسول: ویژگی‌های رمزگذاری شده توسط لایه Bi-GRU به یک شبکه کپسول داده

1. https://nlp.stanford.edu/projects/glove/
2. Hidden vector sequence
3. Schuster and Paliwal

می‌شود. روش‌های مختلفی برای تولید خروجی کپسول وجود دارد که از جمله خروجی را به صورت احتمالاتی تولید می‌کند که این احتمال برابر طول بردار خروجی است. همان‌طور که قبلاً مطرح شد، شبکه کپسول شامل مجموعه‌ای از کپسول‌ها می‌باشد که اندازه آن نشان‌دهنده احتمال وجود ویژگی‌های متناظر است. لایه کپسول ویژگی‌های اسکالر استخراج شده توسط لایه GRU را به کپسول‌های برداری ارزش داده شده تبدیل می‌کند. اگر خروجی Bi-GRU برابر h_i باشد و W ماتریس وزنی باشد، آنگاه $v'_{(i|j)}$ که نشان‌دهنده بردار پیش‌بینی است، از رابطه زیر بدست می‌آید:

$$v'_{(i|j)} = w_{ij}\, h_i$$

مجموعه ورودی‌ها به یک کپسول s_j یک مجموعه وزنی از تمام بردارهای پیش‌بینی $\hat{v}_{(i|j)}$ می‌باشد که از طریق رابطه زیر محاسبه می‌شود:

$$s_j = \sum c_{i,j}.\, v'_{(i|j)}$$

که در این رابطه $c_{i,j}$ ضریب همبستگی می‌باشد که مقدار آن توسط الگوریتم مسیریابی پویا تنظیم می‌شود. در انتها تابع Squash به عنوان یک تابع غیرخطی برای نگاشت مقادیر S_j به بازه [۰-۱] مورد استفاده قرار می‌گیرد. خروجی کپسول در معماری پیشنهادی یک مجموعه بردار می‌باشد. این بدین معنی است که می‌توان انتخاب کرد که خروجی کپسول به کدام کپسول‌ها در لایه بالاتر ارسال شود. در معماری پیشنهادی الگوریتم مسیریابی پویا برای مکانیزم ارسال، مورد استفاده قرار می‌گیرد.

۴) لایه Classification: خروجی شبکه Bi-GRUCapsule پیشنهادی با رویکردهای معمولی متفاوت است، چرا که در این لایه قرار است دو وظیفه متفاوت (تشخیص قطبیت و تشخیص دامنه تعلق) به صورت همزمان انجام گیرند. در ابتدا خروجی هموار شده کپسول که با F نشان داده شده است، به یک لایه تماماً متصل با ۱ نرون داده می‌شود. در این لایه همچنین فراداده‌های مربوط به هر یک از خبرها توسط یک شبکه تماماً متصل یاد گرفته می‌شوند و به عنوان نرون‌های هموار شده به لایه F داده می‌شوند.

$$P = W_{\{dense\}} + W_{\{metadata\}} * F$$

خروجی P باید به گونه‌ای باشد که نشان‌دهنده احتمال هر یک از ۲ کلاس باشد. برای این منظور با استفاده از $Sigmoid$ برای هر $f_i \in F$ مقدار P_i به صورت زیر محاسبه می‌شود:

$$P_i = \frac{1}{1 + e^{-(f_i)}}$$

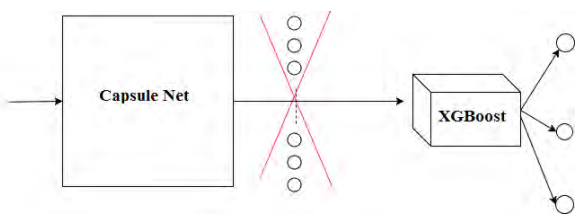
برای تشخیص Fake یا Real بودن، کافی است که $Round(p_i)$ محاسبه شود، اگر مقدار آن ۱ باشد، Real و اگر ۰ باشد Fake است.

۵-۱. اعمال XGBoost

یکی از مهم‌ترین معایب شبکه‌های عصبی عمیق نظیر GRU، CNN، Capsule وجود لایه‌های تماماً متصل می‌باشد که بیشترین تعداد پارامترهای یادگیرنده را شامل می‌شوند. این لایه‌ها وظیفه یادگیری ویژگی‌های استخراج شده توسط لایه‌های میانی را بر عهده دارند. همچنین آخرین لایه از لایه‌های تماماً متصل به عنوان محاسبه‌گر امتیاز دسته‌ها شناخته می‌شود. این لایه توسط یک تابع احتمالاتی نظیر Softmax یا Sigmoid خروجی‌ها را به ازای هر یک از دسته‌ها محاسبه می‌کند. پس از آن توسط یک تابع هزینه میزان خطای شبکه محاسبه می‌شود. جهت کاهش میزان خطای شبکه، خطا به داخل شبکه انتشار داده می‌شود و توسط تابع پس‌انتشار وزن‌های هر یک از نرون‌ها در لایه‌های تماماً متصل و میزان کرنل‌ها در لایه‌های میانی جهت کاهش خطا تغییر داده می‌شود. لایه‌های تماماً متصل در Capsule از نظر محاسباتی سنگین و وقت‌گیر می‌باشند. از طرفی دقت کم‌تری نسبت به طبقه‌بندهایی مانند XGBoost دارند. برای این منظور، طبقه‌بندی ترکیبی را پیشنهاد می‌شود که برای استخراج ویژگی از شبکه‌های Bi-Capsule و برای طبقه‌بندی از XGBoost استفاده می‌کند.

XGBoost توان پیش‌بینی بسیار بالایی دارد که آن را به بهترین گزینه برای دقت در رویدادهای مختلف تبدیل می‌کند؛ چرا که هم مدل خطی و هم الگوریتم یادگیری درختی دارد. این الگوریتم تقریباً ۱۰ برابر سریع‌تر از الگوریتم‌های موجود ارتقای گرادیان است. این الگوریتم شامل تابع‌های عینی مختلف، رگرسیون، کلاس‌بندی و رتبه‌بندی است. یکی از جالب‌ترین نکات در مورد XGBoostاین است که به نام تکنیک ارتقای رگوله شده نیز شناخته می‌شود. این الگوریتم به کاهش مدل‌های بزرگ کمک می‌کند و همچنین پشتیبانی خوبی در طیف وسیعی از زبان‌ها مانند اسکالا، جاوا، R، پایتون، جولیا و C++ دارد.

در شکل شماره (۳) رویکرد پیشنهادی برای طبقه‌بندی اخبار جعل ارائه شده است. در این رویکرد لایه تماماً متصل حذف شده و مدل XGBoost جایگزین آن می‌شود.



شکل ۵- مدل اصلاح شده پیشنهادی برای طبقه‌بندی اخبار جعل

۵-۲. بررسی الگوریتم PSO

PSO یک الگوریتم ازدحام است که از رفتار ذرات / حیوانات اجتماعی مانند ماهی یا پرندگان الهام گرفته شده است. PSO یک روش بهینه‌سازی تصادفی بوده که توسط ابرهارت و کندی (۱۹۹۵)[1] معرفی و توسعه یافته است. همچنین به عنوان یکی از تکنیک‌های فراابتکاری طبقه‌بندی شناخته می‌شود. ایده اصلی الگوریتم PSO ایجاد اشتراک بهتر اطلاعات اجتماعی در میان افراد جامعه است. هر فرد به عنوان یک ذره در ازدحام عمل می‌کند. سپس، آنها یک روش جستجو را در یک فضای جستجو پیاده‌سازی می‌کنند. در طول فرآیند جستجو، آنها اطلاعات و تجربه را برای به روزرسانی مکان‌های بهتر به اشتراک می‌گذارند (نگوین و همکاران[2]، ۲۰۱۹). بنابراین، این روش در جامعه آماری نیز به عنوان یک روش محاسبه تکاملی در نظر گرفته شده است (نگوین و همکاران، ۲۰۱۹؛ چن و همکاران[3]، ۲۰۱۹). الگوریتم PSO پنج مرحله را برای جستجوی بهینه پیاده‌سازی می‌کند:

– **مرحله ۱:** ابتدا جمعیت بومی و سرعت ذرات را مشخص کنید. پس از آن، تناسب ذرات را محاسبه کرده و منطقی‌ترین مکان را به عنوان بهترین محلی و بهترین سراسری[4] تشخیص دهید (لی و همکاران[5]، ۲۰۱۹).

– **مرحله ۲:** هر ذره در فضای جستجو با سرعت اولیه که در مرحله اول مشخص شده است، پرواز می‌کند. سرعت به مقدار بهینه سراسری و محلی بستگی دارد. مطابق فرمول ذیل، سرعت بهینه سراسری و محلی در هر حلقه تکرار، به‌روز می‌شود.

$$v_j^{i+1} = wv_j^{(i)} + \left(c_1 * r_1 * \left(local\ best_j - x_j^{(i)}\right)\right) + \left(c_2 * r_2 * \left(global\ best_j - x_j^{(i)}\right)\right), v_{min} \leq v_{max}$$

که در آن $x_j^{(i)}$ موقعیت ذرات را نشان می‌دهد. $v_j^{(i)}$ نشان‌دهنده سرعت ذره j در تکرار i است. w نشانگر ضریب وزن اینرسی است. i نشان‌دهنده تعداد تکرارها و r_1 و r_2 نماد اعداد در فاصله [۰،۱] است.

1. Eberhart & Kennedy
2. Nguyen & et al.
3. Chen & et al.
4. Global
5. Le & et al.

- مرحله ۳: پس از محاسبه و به‌روزرسانی سرعت جدید، ذرات با سرعت جدید در فضای جستجو پرواز می‌کنند. متناسب با هر موقعیت، سازگاری آنها از طریق عملکرد سازگاری[1] تعیین می‌شود (به عنوان مثال RMSE).

- مرحله ۴: بهترین محلی و سراسری را برای موقعیت بهتر با RMSE پایین به‌روز کنید. بهترین‌های محلی را می‌توان به صورت زیر به‌روز کرد: $x_j^{i+1} = x_j^i + v_j^{(i+1)}; j = 1, 2, \ldots, n$

- مرحله ۵: رضایت جستجو را بررسی کنید. اگر تناسب ذره بهترین (یعنی کم‌ترین RMSE) است، جستجو را متوقف کنید. در غیر این صورت، به مرحله ۲ برگردید.

شبه‌کد الگوریتم PSO برای بهینه‌سازی جستجو در شکل (۴) نشان داده شده است.

Algorithm: The particle swarm optimization (PSO) pseudo-code for the optimization process

```
for each particle i
for each dimention d
Initialize position x_id randomly within permissible range
Initialize velocity v_id randomly within permissible range
end for
end for
Iteration k = 1
do
for each particle i
Calculate fitness value
if the fitness value is better than p_best_id in history
Set current fitness value as the p_best_id
end if
end for
Choose the particle having the best fitness value as the g_best_id
for each particle i
for each dimention d
Calculate velocity according to the following equation
v_j^{i+1} = w v_j^{(i)} + (c1 × r1 × (local best_j − x_j^{(i)})) + (c2 × r2 × (global best_j − x_j^{(i)})), v_min ≤ v_j^{(i)} ≤ v_max
Update particle position according to the following equation
x_j^{i+1} = x_j^{(i)} + v_j^{(i+1)}; j = 1, 2, ..., n
end for
end for
k = k+1
while maximum iterations or minimum error criteria are not attained
```

شکل ٦- فرایند بهینه‌سازی توسط الگوریتم PSO (لی و همکاران، ۲۰۱۹)

1. Fitness function

۶. نتایج آزمایشات

در این بخش پس از معرفی محیط پیاده‌سازی، پارامترهای شبیه‌سازی معرفی خواهد شد. در ادامه نیز با معرفی معیارهای ارزیابی کارایی روش پیشنهادی بر اساس معیارهای ارزیابی موصوف بررسی و یافته‌ها با دیگر روش‌های مشابه مقایسه می‌گردد که برای این مقایسات از رویکرد یادگیری عمیق مبتنی بر شبکه کانولوشن عمیق و دیگر رویکردهای یادگیری عمیق استفاده می‌شود.

۶-۱. محیط پیاده‌سازی

کراس یک کتابخانه سطح بالا برای ایجاد ابزارهای شبکه‌های عصبی بوده که به زبان پایتون نوشته شده است و قابلیت اجرا بر روی TensorFlow, CNTK و Theano را دارد. کراس شامل پیاده‌سازی‌های متعددی از بلوک‌های ساختار شبکه عصبی مانند لایه‌ها، شیءها، توابع فعال‌ساز، بهینه‌سازها و همچنین ابزارهای متعددی برای کار با تصاویر و داده‌های متنی است. این کتابخانه با نسخه‌های 2,7-3,6 پایتون سازگار می‌باشد و می‌تواند به صورت یکپارچه بر روی CPU ها و GPU ها اجرا شود. این API تمرکز خود را بر روی کاربرپسند بودن، ماژولاریتی و توسعه‌پذیری گذاشته است. کراس برای یادگیری و استفاده بسیار آسان می‌باشد، اما این سهولت استفاده باعث کاهش انعطاف‌پذیری آن نمی‌شود. کراس قابلیت ادغام با زبان‌های سطح پایین یادگیری عمیق به‌ویژه تنسورفلو را دارد. در پژوهش حاضر کراس برای ایجاد شبکه‌های عصبی استفاده شد.

۶-۲. معیارهای ارزیابی

با در نظر داشتن پارامترهای زیر می‌توان معیارهای ارزیابی یک طبقه‌بند دودویی را به صورت زیر در نظر گرفت.

- **TP (True Positive):** تعداد نمونه‌های مثبتی که مثبت پیش‌بینی شده‌اند.
- **FP (False positive):** تعداد نمونه‌های منفی است که مثبت پیش‌بینی شده‌اند.
- **TN (True Negative):** تعداد نمونه‌های منفی است که منفی دسته‌بندی شده‌اند.
- **FN (False Negative):** تعداد نمونه‌های مثبتی است که منفی دسته‌بندی شده‌اند.

معیار precision: در یک طبقه‌بند دودویی، precision کسری از اسناد پیش‌بینی شده مثبت که دقیقاً مثبت می‌باشند، نسبت به کل پیش‌بینی‌های مثبت انجام شده، می‌باشد (وکیلی و همکاران[1]، ۲۰۲۰).

1. Vakili & et al.

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

به طور کلی Precision بالا به معنی توانایی بالا برای پیش‌بینی درست است.

معیار Recall: نسبت تعداد اسناد پیش‌بینی‌شده مثبت نسبت به کل اسناد موجود شامل پیش‌بینی مثبت و منفی می‌باشد (وکیلی و همکاران، ۲۰۲۰).

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

Recall بالا به معنی این است که طبقه‌بند توانسته میزان زیادی از اسناد را درست طبقه‌بندی کند.

F-Measure: میانگین هارمونیک نیز نامیده می‌شود که ترکیبی از دو معیار Precision و Recall می‌باشد. معمولاً به آن F-Score متعادل نیز می‌گویند (وکیلی و همکاران، ۲۰۲۰).

$$F_{score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Accuracy: یکی دیگر از معیارهایی است که در یک طبقه‌بند دودویی مورد استفاده قرار می‌گیرد. برای محاسبه آن نسبت TP وTN را به تعداد کل موارد پیش‌بینی شده در نظر می‌گیرند (وکیلی و همکاران، ۲۰۲۰).

$$Accuracy = \frac{TP + FN}{TP + FP + FN + TN}$$

۶-۳. بررسی نتایج حاصل از اجرای رویکرد پیشنهادی کپسول

جدول شماره (۳) نشان‌دهنده دقت بدست آمده از رویکرد پیشنهادی کپسول بر روی مجموعه داده‌های مورد نظر می‌باشد. این رویکرد با ۵ رویکرد متفاوت دیگر که در پژوهش یانگ و همکاران[1] (۲۰۱۸) معرفی شده‌اند، مورد مقایسه قرار گرفت. این رویکردها عبارتند از:

۱) مدل Character base CNN: مدل دنباله‌ای از کاراکترهای رمزگذاری شده را به عنوان ورودی می‌پذیرند. رمزگذاری با تجویز الفبای اندازه m برای زبان ورودی انجام می‌شود و سپس با استفاده از رمزگذاری m−۱ (یا رمزگذاری "one-hot") هر کاراکتر را به مجموعه‌ای از بردارها تبدیل می‌کنیم. سپس، دنباله کاراکترها به دنباله‌ای از چنین بردارهایی با اندازه m با طول ثابت تبدیل می‌شود. الفبای مورد استفاده در تمام مدل‌ها شامل ۷۶ کاراکتر است که شامل ۳۲ حرف فارسی،

1. Yang & et al.

۱۰ رقم، ۳۳ کاراکتر دیگر که علائم نگارشی و... می‌باشند. این مدل جهت مقایسه با مدل‌های پیشنهادی مورد استفاده قرار گرفت. ساختار این مدل در پژوهش کاچکینا و همکاران (۲۰۱۸) آمده است.

۲) مدل Multi-Channel CNN: در پژوهش کیم[1] (۲۰۱۴)، یک شبکه CNN ساده با یک لایه کانولوشن در بالای بردارهای کلمه بدست آمده از یک مدل زبانی غیرنظارتی آموزش داده شده است. این بردارها توسط Mikolov و همکاران (۲۰۱۳) در ۱۰۰ میلیارد کلمه گوگل منتشر شده‌اند و در دسترس عموم قرار دارند. این معماری به عنوان یکی دیگر از معماری‌ها جهت مقایسه با رویکردهای پیشنهادی مورد استفاده قرار گرفت.

۳) ماشین‌های بردار پشتیبان: یکی از متداول‌ترین روش‌های طبقه‌بندی در برخی از زمینه‌های تحقیقاتی است. SVM ها طبقه‌بندهای تبعیض‌آمیز هستند که به طور رسمی توسط یک مرز جداکننده تعریف می‌شوند و سعی در یافتن مرزی دارند که بیشترین حاشیه امن را بین مجموعه داده‌ها داشته باشد.

۴) درخت تصمیم‌گیری: یکی دیگر از خانواده‌های الگوریتم‌هایی که به طور گسترده، به ویژه برای کارهای تجزیه و تحلیل شایعه مورد مطالعه قرار گرفته‌اند، درخت تصمیم‌گیری است (ژو و همکاران، ۲۰۰۳). درخت تصمیم به منظور تعیین کلاس، تقسیم بازگشتی روی مقادیر ویژگی‌ها انجام می‌دهد. درخت تصمیم‌گیری برای داده‌هایی توسط الگوریتم‌هایی مانند J48 (C4.5) (چانگ و همکاران، ۲۰۱۶) تولید می‌شود. علی‌رغم سادگی نسبی آنها نسبت به سایر طرح‌های یادگیری ماشین، آنها عملکرد رقابتی را در بسیاری از وظایف بدست آورده‌اند.

۵) جنگل تصادفی: جنگل‌های تصادفی (ژو و همکاران، ۲۰۰۳) در تعدادی از کارهای تحلیل شایعه به‌کار گرفته شده است. یک جنگل تصادفی مجموعه‌ای از چندین درخت تصمیم است که با ترکیب پیش‌بینی‌ها از هر درخت، خروجی نهایی را بدست می‌آورد. مطالعات تطبیقی در مورد الگوریتم‌های مختلف یادگیری ماشین برای شایعات و کارهای خبری جعلی، نشان می‌دهد که جنگل‌های تصادفی می‌تواند به عنوان یک طبقه‌بند قوی برای تشخیص مورد استفاده قرار گیرد (آکر و همکاران، ۲۰۱۷).

1. Kim

رویکرد پیشنهادی به دلیل در نظر گرفتن موقعیت جمله‌ها در متن در سه معیار ارزیابی بررسی شده، دقت بالاتری را بدست آورده است.

جدول ۳- بررسی نتایج بدست آمده توسط رویکرد پیشنهادی و سایر رویکردها

Methods	Accuracy	Precision	Recall	F1
Decision Tree	0.9200	0.9200	0.9200	0.9200
Random Forest	0.9100	0.9100	0.9100	0.9100
SVM	0.9100	0.9300	0.9100	0.9200
Char based model	0.8400	0.9300	0.7700	0.8700
Multi-channel CNN	0.8800	0.8800	0.8800	0.8800
Bi-GRUCapsule	0.9461	0.9490	0.9433	0.9593

در پاسخ به سؤال شماره (۱) تحقیق: طبق جدول شماره (۳)، سه رویکرد یادگیری ماشین (Random Forest[۱], Decision Tree[۲] و SVM[۳]) دقت تقریباً برابری را بدست آورده‌اند. این در حالی است که رویکردهای مبتنی بر شبکه‌های کانولوشن معمولی مانند Character based model و Multi-Channel CNN دقت کم‌تری نسبت به آن‌ها داشته‌اند. این میزان اختلاف ناشی از عدم استخراج ویژگی‌های مناسب برای رویکردهای یادگیری عمیق کانولوشن می‌باشد. در واقع این رویکردها نتوانسته‌اند ویژگی‌های مناسب را استخراج کنند. از طرف دیگر این اختلاف نشان‌دهنده قدرت رویکردهای یادگیری ماشین می‌باشد که می‌توانند بر روی مسائل طبقه‌بندی به دقت مناسبی برسند.

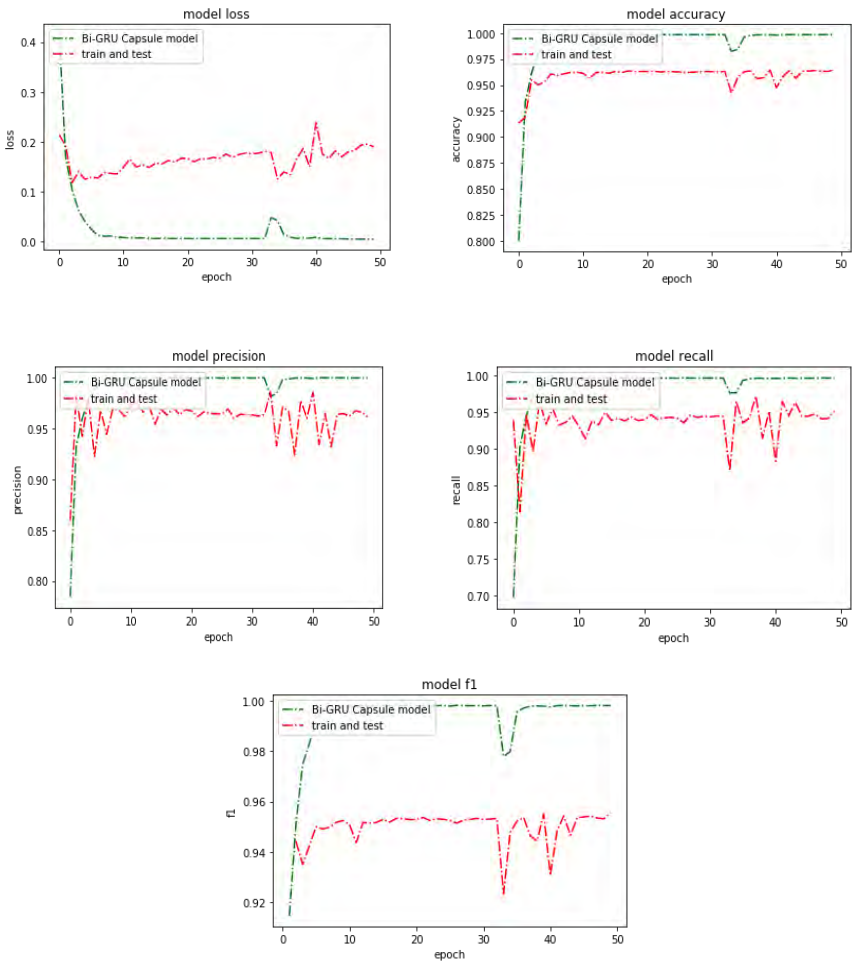
در پاسخ به سؤال (۲) تحقیق، رویکرد پیشنهادی Bi-GRUCapsule با توجه به در نظر گرفتن مکان قرارگیری جملات توانسته نسبت به سایر رویکردها به دقت بهتری دست یابد. تغییرات انجام شده در لایه‌های شبکه‌های کانولوشن باعث شده است شبکه کانولوشن بتواند ویژگی‌های مناسبی را استخراج کند و نسبت به رویکردهای شبکه‌های کانولوشن معمولی و رویکردهای یادگیری ماشین معمولی به دقت بهتری دست یابد.

یکی از مهم‌ترین پارامترها و معیارها برای مقایسه F1 می‌باشد که رویکرد پیشنهادی توانست به ۰/۹۵۹۳ برسد و نسبت به بهترین رویکرد موجود که ۰/۹۲ می‌باشد، حدوداً ۴ درصد بهبود پیدا

۱. جنگل تصادفی
۲. درخت تصمیم
۳. ماشین بردار پشتیبان

کند که این به نوبه خود قابل توجه است.

نمودار (۷) مربوط به میزان خطای مدل پیشنهادی برای ۵۰ مرحله به اجرا درآمد. همچنین نمودارهای مربوط به Accuracy، Precision، Recall و F1 رویکرد پیشنهادی در ادامه آمده است.



شکل ۷- نمودارهای مربوط به Accuracy، Precision، Recall و F1 رویکرد پیشنهادی اول

۶-۴. بررسی نتایج حاصل از اجرای رویکرد پیشنهادی Bi-GRU Capsule+XGBoost

برای این رویکرد مقادیر پارامترهای زیادی مورد تست و ارزیابی قرار گرفت. با توجه به جدول (۴)، دقت‌های بدست آمده توسط رویکرد پیشنهادی با مقادیر Batch Size متفاوت است.

جدول ٤- مقادیر مختلف بدست آمده توسط رویکرد پیشنهادی با مقادیر اندازه دسته متفاوت

Batch size	Precision	Recall	F1
8	0.8723	0.8756	0.8730
16	0.8744	0.8846	0.8813
128	0.8912	0.8990	0.8959
256	0.9146	0.8704	0.8920
512	0.9273	0.9688	0.9466

برای این رویکرد همچنین مقادیر متفاوتی از لایه‌های تماماً متصل نیز مورد تست قرار گرفت. نتایج بدست آمده در جدول شماره (۵) نشان‌دهنده این مقادیر می‌باشد:

جدول ٥- مقادیر مختلف بدست آمده توسط رویکرد پیشنهادی با مقادیر لایه تماماً متصل متفاوت

Batch size	Precision	Recall	F1
8	0.8733	0.8716	0.8726
16	0.8744	0.8850	0.8846
128	0.8913	0.9000	0.8960
256	0.9273	0.9688	0.9466
512	0.9246	0.9404	0.9220

نتایج نشان‌دهنده این است که انتخاب مقادیر بهینه در دقت نهایی تأثیرگذار می‌باشد. برای این منظور الگوریتم PSO را برای یافتن مقادیر بهینه بر روی طبقه‌بند اعمال می‌کنیم. این رویکرد نسبت به رویکرد پیشنهادی دقت بالاتری را بدست آورد. همچنین در مقایسه با ۶ رویکرد بررسی شده قبلی نیز دقت بالاتری را بدست آورد. جدول شماره (۶) نتایج حاصل از این رویکرد بر روی مجموعه داده‌های بررسی شده می‌باشد.

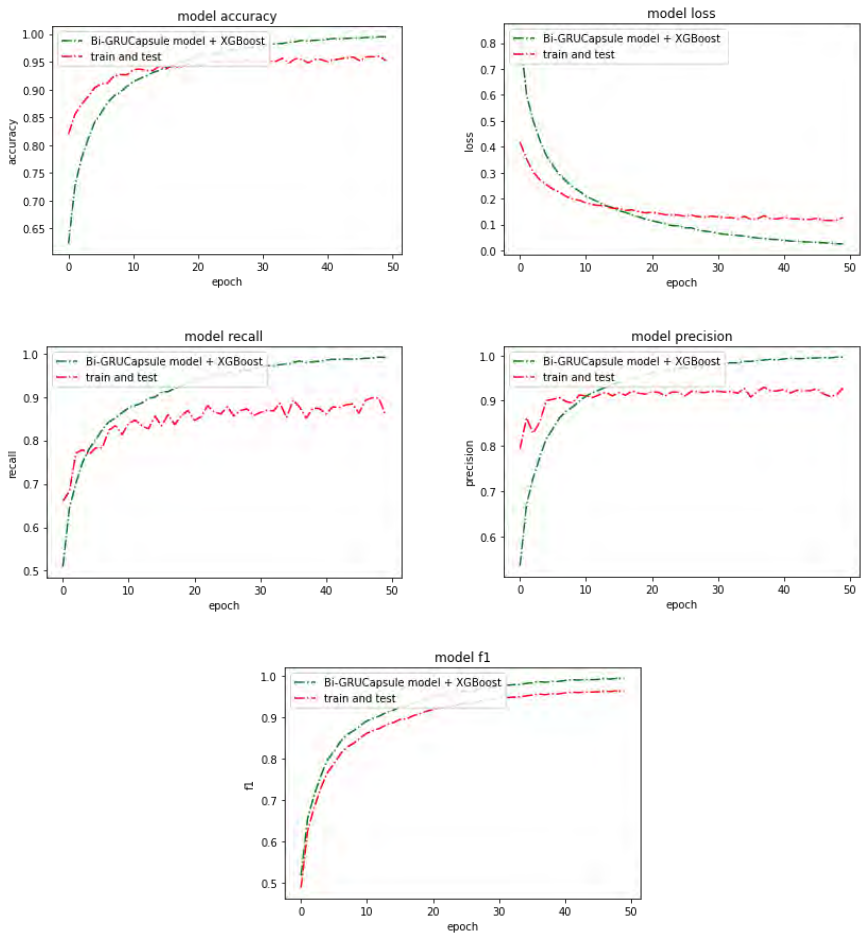
طبق سؤال ۳ تحقیق این رویکرد نسبت به رویکرد اول که دارای لایه‌های تماماً متصل بود، دقت بالاتری دارد. دلیل این امر استفاده از XGBoost به عنوان طبقه‌بند در لایه آخر است. هر یک از این رویکردها دارای پارامترهایی هستند که با توجه به پیوسته بودن بازه انتخاب این مقادیر نیازمند الگوریتم‌های جستجو برای یافتن مقادیر بهینه است.

جدول ٦- نتایج بدست آمده توسط رویکرد پیشنهادی دوم

Methods	Accuracy	Precision	Recall	F1
Decision Tree	0.9200	0.9200	0.9200	0.9200
Random Forest	0.9100	0.9100	0.9100	0.9100
SVM	0.9100	0.9300	0.9100	0.9200
Char based model	0.8400	0.9300	0.7700	0.8700

Methods	Accuracy	Precision	Recall	F1
Multi-channel CNN	0.8800	0.8800	0.8800	0.8800
Bi-GRUCapsule	0.9461	0.9490	0.9433	0.9593
Bi-GRUCapsule+XGBoost+PSO	0.9555	0.9591	0.9538	0.9595

شکل (۸) مربوط به میزان خطای مدل پیشنهادی برای ۵۰ مرحله است. همچنین نمودارهای مربوط به Accuracy، Precision، Recall و F1 رویکرد پیشنهادی در ادامه آمده‌اند.



شکل ۸ – نمودارهای مربوط به Accuracy، Precision، Recall و F1 رویکرد پیشنهادی دوم

در پاسخ به سؤال ۳ تحقیق: بررسی نمودارهای رویکرد دوم گویای این است که خطا به صورت لگاریتمی پایین آمده و هیچ‌گونه پیش‌برازشی اتفاق نیافته است. همچنین در مراحل اجرای بالاتر

امکان بدست آمدن دقت بالاتر امکان‌پذیر است. اما برای مقایسه در شرایط یکسان، مقدار ۵۰ در نظر گرفته شد.

۷. نوآوری پژوهش

هدف تحقیق حاضر ارائه رویکرد ترکیبی برای اخبار جعل می‌باشد که از ترکیب شبکه Bi-GRUCapsule و XGBoost برای این منظور استفاده می‌شود. جنبه‌های نوآوری این تحقیق عبارتند از:

۱) استفاده از شبکه Bi-GRUCapsule برای طبقه‌بندی اخبار فارسی،

۲) جمع‌آوری مجموعه داده‌های فارسی اخبار جعل،

۳) استفاده از ترکیب XGBoost و رویکردهای یادگیری عمیق،

۴) بهینه‌سازی پارامترهای شبکه با PSO.

۸. نتیجه‌گیری

در این پژوهش رویکردی مبتنی بر شبکه کپسول برای طبقه‌بندی اخبار جعل مورد بررسی قرار گرفت. همچنین به نتایج حاصل از رویکرد پیشنهادی Bi-GRUCapsule بر روی مجموعه داده‌های مورد نظر با توجه به معیارهای ارزیابی پرداخته شد. برای حل مشکل وجود لایه‌های تماماً متصل از رویکرد XGBoost استفاده شده و برای فضای پارامتری، الگوریتم PSO بر روی آن اعمال شد. نتایج نشان می‌دهد که در حالت کلی رویکرد پیشنهادی عملکرد بهتری را نسبت تمامی رویکردهای پیشنهادی قبلی ازجمله الگوریتم‌های پایه و الگوریتم‌های یادگیری ماشین داشته و به ۹۶ درصد دقت و صحّت دست یافته و با توجه به اینکه داده‌های مربوط به اخبار فارسی کرونا بسیار کم بوده، طی بررسی مشخص شد که بهبود الگوریتم‌های یادگیری عمیق در مورد تشخیص اخبار جعل فارسی به شدت کم بوده و شاید نادر می‌باشد. این رویکرد با پارامترهای متفاوتی مورد ارزیابی و تست قرار گرفت که نتایج موجود در جداول بررسی شده بهترین مقدار بدست آمده با این پارامترها بوده است. یکی از مهم‌ترین مسائلی که در طبقه‌بندی اخبار جعل در نظر گرفته نمی‌شود، بررسی حوزه نفی می‌باشد. در واقع حوزه کلمه نفی، تأثیر یک کلمه منفی را نشان می‌دهد که تا کجای جمله می‌تواند تأثیر بگذارد. رویکردهای متفاوتی برای این منظور در ادبیات پیشنهاد شده است. برای این منظور برای کارهای آینده می‌توان از الگوریتم حوزه نفی به همراه الگوریتم پیشنهادی برای بهبود طبقه‌بند موردنظر استفاده کرد.

References

Aker, A., Derczynski, L. & Bontcheva, K. (2017). *Simple open stance classification for rumour analysis.* Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017, (pp. 31–39). **DOI**: 10.26615/978-954-452-049-6_005

Allcott, H. & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2): 211-236. **DOI:** 10.1257/jep.31.2.211

Allport, G.W. & Postman, L. (1946). An analysis of rumor. *Public opinion quarterly*, 10(4): 501-517. **DOI**: 10.1086/265813

Briscoe, E.J., Appling, D.S. & Hayes, H. (2014). *Cues to deception in social media communications*. In: 2014 47th Hawaii international conference on system sciences (pp. 1435-1443). IEEE. **DOI:** 10.1109/HICSS.2014.186

Bondielli, A. & Marcelloni, F. (2019). A survey on fake news and rumour detection techniques. *Information Sciences*, 497: 38-55. **DOI:** 10.1016/j.ins.2019.05.035

Castillo, C., Mendoza, M. & Poblete, B. (2011). *Information credibility on twitter*. In: Proceedings of the 20th international conference on World wide web (pp. 675-684). **DOI:** 10.1145/1963405.1963500

Chang, C., Zhang, Y., Szabo, C. & Sheng, Q.Z. (2016). *Extreme user and political rumor detection on twitter*. In: International conference on advanced data mining and applications (pp. 751-763). Springer, Cham.

Chen, H., Asteris, P.G., Jahed Armaghani, D., Gordan, B. & Pham, B.T. (2019). Assessing dynamic conditions of the retaining wall: developing two hybrid intelligent models. *Applied Sciences*, 9(6):1042. **DOI:** 10.3390/app9061042

Chen, Y.C., Liu, Z.Y. & Kao, H.Y. (2017). Ikm at semeval-2017 task 8: Convolutional neural networks for stance detection and rumor verification. In: *Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017)* (pp. 465-469). **DOI**: 10.18653/v1/S17-2081

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. & Bengio, Y. (2014). *Learning phrase representations using RNN encoder-decoder for statistical machine translation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).* Association for Computational Linguistics, pp.1724-1734. **DOI**: 10.3115/v1/D14-1179

Conroy, N.K., Rubin, V.L. & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the association for information science and technology*, 52(1), P.1-4. **DOI:** 10.1002/pra2.2015.145052010082

Eberhart, R. & Kennedy, J. (1995). *A new optimizer using particle swarm theory*. In: MHS'95. Proceedings of the sixth international symposium on micro machine and human science (pp. 39-43). Ieee. **DOI:** 10.1109/MHS.1995.494215

Giasemidis, G., Singleton, C., Agrafiotis, I., Nurse, J.R., Pilgrim, A., Willis, C. & Greetham, D.V. (2016). *Determining the veracity of rumours on Twitter*. In: International Conference on Social Informatics (pp. 185-205). Springer, Cham. **DOI**: 10.1007/978-3-319-47880-7_12

Gorrell, G., Bontcheva, K., Derczynski, L., Kochkina, E., Liakata, M. & Zubiaga, A. (2018).

Rumoureval 2019: Determining rumour veracity and support for rumours. Proceedings of the 13th International Workshop on Semantic Evaluation: 845–854. **DOI:** 10.18653/v1/S19-2147

Jacovi, A., Shalom, O.S. & Goldberg, Y. (2018). *Understanding convolutional neural networks for text classification.* Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP. Association for Computational Linguistics: 56-65. **DOI**: 10.18653/v1/W18-5408

Kim,Y. (2014). *Convolutional Neural Networks for Sentence Classification*. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics: 1746–1751. **DOI:** 10.3115/v1/D14-1181

Kochkina, E., Liakata, M. & Zubiaga, A. (2018). *All-in-one: Multi-task learning for rumour verification.* Association for Computational Linguistics. Proceedings of the 27th International Conference on Computational Linguistics: 3402–3413.

Kwon, S., Cha, M., Jung, K., Chen, W. & Wang, Y. (2013). *Prominent features of rumor propagation in online social media*. In: 2013 IEEE 13th international conference on data mining (pp. 1103-1108). IEEE. **DOI:** 10.1109/ICDM.2013.61

Le, L.T., Nguyen, H., Zhou, J., Dou, J. & Moayedi, H. (2019). Estimating the heating load of buildings for smart city planning using a novel artificial intelligence technique PSO-XGBoost. *Applied Sciences*, 9(13): 2714**. DOI:** 10.3390/app9132714.

Le, L.T., Nguyen, H., Dou, J. & Zhou, J. (2019). A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning. *Applied Sciences*,9(13): 2630. **DOI:** 10.3390/app9132630.

LeCun, Y., Kavukcuoglu, K. & Farabet, C. (2010). *Convolutional networks and applications in vision*. In: Proceedings of 2010 IEEE international symposium on circuits and systems (pp. 253-256). IEEE. **DOI:** 10.1109/ISCAS.2010.5537907

Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B.J., Wong, K.F. & Cha, M. (2016). *Detecting rumors from microblogs with recurrent neural networks*. Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016): 3818-3824.

Meyer, J.K. (1969). *Bibliography on the urban crisis: The behavioral, psychological, and sociological aspects of the urban crisis (no. 1948).* National Institute of Mental Health.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26: 3111–3119.

Miller, T., Howe, P. & Sonenberg, L. (2017). *Explainable AI: Beware of inmates running the asylum or: How I learnt to stop worrying and love the social and behavioural sciences.* arXiv preprint arXiv: 1712.00547. **DOI:** https://doi.org/10.48550/arXiv.1712.00547

Moon, A. (2017). *Two-thirds of American adults get news from social media: survey*. Recuperado de: https://uk. reuters. com/article/us-usa-internet-socialmedia/two-thirds-of-american-adults-get-news-from-social-media-survey-idUKKCN1BJ2A8.

Nguyen, H.H., Yamagishi, J. & Echizen, I. (2019). Use of a capsule network to detect fake images and videos arXiv 2019. *arXiv preprint arXiv:1910.12467*. **DOI:** 10.48550/arXiv.1910.12467

Qin, Y., Wurzer, D., Lavrenko, V. & Tang, C. (2016). *Spotting rumors via novelty detection*. **DOI:** 10.48550/arXiv.1611.06322.

Rathnayaka, P., Abeysinghe, S., Samarajeewa, C., Manchanayake, I. & Walpola, M. (2018). *Sentylic at IEST 2018: Gated recurrent neural network and capsule network based approach for implicit emotion detection. Association for Computational Linguistics*. Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis: p.254-259. **DOI:** 10.18653/v1/W18-6237

Rubin, V.L., Chen, Y. & Conroy, N.K. (2015). *Deception detection for news: three types of fakes*. Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community, 2015, p. 83: American Society for Information Science. **DOI:** 10.1002/pra2.2015.145052010083.

Rubin, V.L., Conroy, N., Chen, Y. & Cornwell, S. (2016). *Fake news or truth? using satirical cues to detect potentially misleading news*. In: Proceedings of the second workshop on computational approaches to deception detection (pp. 7-17). **DOI:** 10.18653/v1/W16-0802

Ruchansky, N., Seo, S. & Liu, Y. (2017). *Csi: A hybrid deep model for fake news detection.* In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (pp. 797-806). **DOI:** 10.1145/3132847.3132877

Sabour, S., Frosst, N. & Hinton, G.E. (2017). *Dynamic routing between capsules.* Advances in neural information processing systems, Proceedings of the 31st International Conference on Neural Information Processing Systems: 3859-3869.

Schuster, M. & Paliwal, K.K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11): 2673-2681. **DOI:** 10.1109/78.650093

Vakili, M., Ghamsari, M. & Rezaei, M. (2020). *Performance analysis and comparison of machine and deep learning algorithms for IoT data classification.* Third International Conference on Computing and Network Communications. **DOI:** 10.48550/arXiv.2001.09636

Vosoughi, S. (2015). *Automatic detection and verification of rumors on Twitter*.Doctoral dissertation, Massachusetts Institute of Technology. Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Arts and Sciences.

Vosoughi, S., Mohsenvand, M.N. & Roy, D. (2017). Rumor gauge: Predicting the veracity of rumors on Twitter. *ACM transactions on knowledge discovery from data (TKDD)*, 11(4): 1-36. **DOI:** 10.1145/3070644

Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z. & Yu, P.S. (2018). TI-CNN: Convolutional neural networks for fake news detection. *arXiv preprint arXiv:1806.00749*. **DOI:** 10.48550/arXiv.1806.00749

Yang, F., Liu, Y., Yu, X. & Yang, M. (2012). *Automatic detection of rumor on sina weibo*. In: Proceedings of the ACM SIGKDD workshop on mining data semantics (pp. 1-7). **DOI:** 10.1145/2350190.2350203

Zeng, L., Starbird, K. & Spiro, E.S. (2016). *# unconfirmed: Classifying rumor stance in crisis-related social media messages*. In: Tenth International AAAI Conference on Web and Social Media.

Zhang, H., Fan, Z., Zheng, J. & Liu, Q. (2012). An improving deception detection method in

computer-mediated communication. *Journal of Networks*, 7(11): 1811. **DOI**:10.4304/jnw.7.11.1811-1816

Zhou, L., Twitchell, D.P., Qin, T., Burgoon, J.K. & Nunamaker, J.F. (2003). *An exploratory study into deception detection in text-based computer-mediated communication*. In: 36th Annual Hawaii JanuaryInternational Conference on System Sciences, 2003. IEEE. **DOI:** 10.1109/HICSS.2003.1173793

Zubiaga, A., Liakata, M. & Procter, R. (2016). Learning reporting dynamics during breaking news for rumour detection in social media. *arXiv preprint arXiv:1610.07363*. **DOI:** 10.48550/arXiv.1610.07363

Zubiaga, A., Liakata, M., Procter, R., Wong Sak Hoi, G. & Tolmie, P. (2016). Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PloS one*, 11(3): e0150989. **DOI:** doi.org/10.1371/journal.pone.0150989

Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M. & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys (CSUR)*, 51(2): 1-36. **DOI:** 10.1145/3161603