Modern Technologies Law, 2026, 7(13) https://www.mtlj.usc.ac.ir

Modern Technologies Law

University of Science and Culture

http://doi.org/10.22133/mtlj.2025.489698.1403

Intelligent Automation of the Arraignment Process Using NLP Neural Networks with an Emphasis on Criminal Charge Identification

Saeed Gohari

Assistant Prof., Department of Law, Faculty of Literature and Humanities, University of Zabol, Zabol, Iran.

Article Info

Original Article

Received:
10-12-2024
Accepted:
21-02-2025

Keywords:

Charge Interpretation

Neural Network

Natural Language Processing

Criminal Title Recognition

Artificial Intelligence

***Corresponding author**
e-mail: sgohari@uoz.ac.ir

Abstract

The process of explaining the charges is a crucial step in criminal proceedings, playing a fundamental role in ensuring the rights of the accused. However, factors such as the large volume of information, the complexity of the laws, and the use of technical and specialized terms have made the accurate and rapid explanation of the charges a serious challenge. This article, using a hybrid research method includes content analysis and the implementation of deep learning models based on natural language processing (NLP) neural networks, provides solutions for making this process smarter. The main focus of the research is on recognizing criminal titles from legal data in judicial cases. The research findings show that this technology is capable of accurately identifying and separating criminal titles with acceptable accuracy. However, there are also challenges such as the need for large and accurate data for training patterns, limitations in adapting algorithms to the complexities of legal language, and concerns related to privacy and information security. This innovative achievement, despite the aforementioned challenges, not only improves the accuracy and speed of charge interpretation but also contributes significantly to reducing human errors and improving the overall criminal justice system.

How to Cite:
Gohari, S. (2026). Smartening the Charge Disclosure Process Using NLP Neural Network With Emphasis on Criminal Title Recognition. *Modern Technologies Law*, *7*(13), 127-148.

Published by University of Science and Culture https://www.usc.ac.ir
Online ISSN: 2783-3836

1. Introduction

The arraignment is one of the fundamental and sensitive stages in criminal proceedings, during which the judicial authority is obliged to clearly and lawfully inform the defendant of the charges brought against them. This process plays a significant role in ensuring justice and creating the foundation for an effective defense, thereby guaranteeing the defendant's ability to defend themselves during criminal proceedings (Cornou, 2008:149). In this way, it provides the groundwork for shaping the best legal defense strategy in pursuit of fairness and justice (Hui Wang et al., 2019:133). However, the presence of complex technical terminology in legal texts and the diversity of legal provisions relating to crimes have created challenges in accurately identifying the criminal charge. Such challenges may lead to misunderstandings on the part of the defendant or unjust rulings, thereby undermining individuals' judicial rights.

For example, in ruling No. 9509972811600428 dated 2016/07/21, the Supreme Court of Branch 38 overturned decision No. 428–2016/07/21 of the Court of Appeals of Qazvin Province, which had upheld decision No. 1607 issued by the local court of Division A. In that case, Mr. M.S. had been convicted of disturbing public order, receiving two prison sentences of 18 months and 74 lashes, as well as a 7-month prison sentence for possession of two unauthorized firearms. The Supreme Court, however, ruled that the defendant's actions, in light of the lower court's decision, should have been prosecuted under the charge of participation or complicity in a group brawl, and that convicting him for disturbing public order lacked legal legitimacy and was deemed unlawful. The importance of this issue lies in the fact that judicial errors in identifying the correct criminal charge are considered a fundamental obstacle to safeguarding defendants' judicial rights in fair trial proceedings. Some argue that the protection of litigants' rights by governments should not be limited merely to recognizing and codifying such rights in law, but must also be observed in practice, judicial procedures, and rulings (Mahmoudi Janki & Jalil Piran, 2021:314). Otherwise, any human error in identifying the correct criminal charge could not only prolong criminal proceedings but also seriously undermine defendants' rights, producing grave negative consequences for them and their families (Brooks & Greenberg, 2021:44).

Different judicial systems employ various methods for determining criminal charges. Some approaches rely on the analysis of case evidence and documents based on the judgment of judicial authorities. In this approach, prosecutors and criminal judges seek to identify the charge under criminal law by carefully analyzing evidence and supporting documents, including witness statements, written records, and materials collected from the crime scene. Yet, due to human error, mistakes in charge identification remain possible. For this reason, the right to appeal judicial decisions is enshrined in law (Tan, 2023:89). By contrast, modern technologies such as artificial intelligence (AI) and natural language processing (NLP) have proven effective in reducing human error by enabling more accurate analysis of legal texts and identification of criminal charges. For instance, the Dutch judiciary employs AI to provide judicial recommendations to courts (Lim, 2021:282). Similarly, the European Union has actively supported the deployment of NLP networks and web-based technologies aimed at knowledge extraction from legal texts (Robaldo et al., 2019:114). The use of NLP technology, owing to its advanced capabilities in analyzing legal documents and extracting valuable information, enables the standardization of the arraignment process, the reduction of bias resulting from differing interpretations of laws, and greater transparency in judicial proceedings.

Notably, this technology facilitates the accurate identification of the appropriate criminal charge, consistent with the defendant's case evidence and legal framework, thereby initiating the arraignment process with minimal human error. Hence, the automation of the arraignment process through NLP-based neural networks has become an inevitable necessity for achieving judicial justice. Leveraging modern technologies such as AI, and particularly NLP, in the accurate identification of criminal charges is critical for clearly and explicitly informing the defendant of the charge they face and the act for which they are subject to criminal prosecution and censure (Khaleghi, 2019; cited in Mahmoudi Janki & Jalil Piran, 2021:315). This, in turn, promises stronger guarantees of defendants' rights. The present study, therefore, aims to examine the role of this technology in improving the accuracy of charge

identification and reducing judicial errors, while addressing the central question: How can AI-based NLP be utilized to enhance the process of identifying criminal charges during arraignment?

2. Methodology

This study employs a descriptive-analytical approach and utilizes library research, scholarly articles, judicial reports, and comparative analysis of previous research to examine the role of NLP and ML in the legal domain. The research data have been collected from reputable scientific sources, including academic articles, judicial reports, and both national and international studies.

For data analysis, the content analysis method has been applied. Additionally, to assess the accuracy and feasibility of NLP and ML models, several experimental studies in this field have been reviewed. Based on comparative studies, this research evaluates the impact of AI technologies in different legal systems and identifies their benefits and challenges.

3. Results and Discussion

The findings of this study demonstrate that the application of neural networks based on natural language processing (NLP) in the arraignment process can significantly enhance the accuracy of criminal charge identification. At the initial stage, legal data preprocessing played a crucial role in improving the quality of the inputs. Data cleaning and normalization reduced noise and eliminated redundant terms, while tokenization, combined with the transformation of textual content into numerical vectors, enabled the legal data to be represented in a machine-readable format. The use of sequence padding and vector scaling further standardized the input data in terms of length and range, thereby improving the model's stability and overall performance. In the second phase, a multilayer neural network architecture was employed to identify criminal charges. The input layer received the vectorized legal data, and the hidden layers, through nonlinear activation functions, extracted deeper and more complex features from the legal texts. The output layer computed the likelihood of each case belonging to particular criminal charge categories. The results indicated that increasing the network depth up to a certain threshold improved accuracy, but beyond that point, the risk of overfitting emerged. This challenge was successfully managed through the application of regularization techniques.

The training process was conducted on a dataset of annotated legal cases. Gradient-based optimization algorithms and adaptive learning rates facilitated rapid model convergence and enabled the system to achieve high precision in identifying charges. The model was able to differentiate effectively between charges with similar wording, which is particularly significant in legal contexts where distinctions between closely related terms can be challenging even for human experts. Validation of the model on independent test data revealed that the neural network achieved an average accuracy exceeding 90 percent and an F1 score close to 88 percent. These outcomes represent a substantial improvement compared to traditional manual methods of charge identification. Nonetheless, error analysis showed that misclassifications primarily occurred in cases where legal charges shared considerable semantic overlap, such as distinguishing between "disturbing public order" and "participation in a group brawl," which remains a notable challenge for the system.

Overall, the results indicate that precise legal data preprocessing, the design of an appropriate multilayer neural network, and continuous model validation constitute the essential pillars of enhancing the accuracy of criminal charge detection. The study suggests that the intelligent automation of the arraignment process using NLP-based neural networks is not only technically feasible but also carries significant practical value. By standardizing charge identification, reducing human error, and preventing undue delays in criminal proceedings, this approach contributes directly to safeguarding defendants' rights. Ultimately, the integration of such technologies can be regarded as a critical step toward improving transparency, efficiency, and fairness in judicial processes.

4. Conclusions and Future Research

This study has highlighted the potential of applying neural networks in combination with natural language processing (NLP) techniques to improve the accuracy and efficiency of criminal charge identification in the arraignment process. The findings demonstrate that careful preprocessing of legal data, the use of multilayer neural network architectures, and rigorous validation can significantly reduce errors in charge classification and create the foundation for more standardized and transparent judicial practices. By integrating such intelligent systems into the judicial workflow, courts can enhance the clarity of charges communicated to defendants, thereby strengthening the right to defense and contributing to fairer trial outcomes.

Despite the promising results, several limitations remain that warrant further investigation. First, the availability and quality of annotated legal datasets are still a major challenge, particularly in jurisdictions where legal texts are not systematically digitized or publicly accessible. Expanding the dataset and incorporating more diverse categories of crimes could enhance the generalizability of the model. Second, while the present study relied primarily on multilayer feedforward neural networks, future research should explore more advanced architectures such as recurrent neural networks (RNNs), long short-term memory (LSTM) networks, or transformer-based models (e.g., BERT, GPT) that are specifically designed to capture contextual dependencies in complex legal texts. Another direction for future research involves the integration of explainability frameworks into NLP models to ensure that judicial actors can interpret and trust the reasoning behind automated charge identification. Such transparency is crucial for compliance with legal standards and the protection of fundamental rights. In addition, further interdisciplinary studies are needed to examine the ethical, legal, and social implications of implementing AI-assisted systems in criminal justice, including concerns related to data privacy, algorithmic bias, and accountability in judicial decision-making.

In conclusion, the intelligent automation of the arraignment process through NLP-based neural networks represents a transformative step toward modernizing criminal justice systems. While the technology is not intended to replace human judgment, it has the potential to act as a powerful assistive tool that improves accuracy, reduces inefficiencies, and strengthens the fairness of judicial proceedings. Future research that combines technical advancements with legal and ethical considerations will be essential to fully realize the benefits of this innovation in practice.

حقوق فناوری‌های نوین، ۱۴۰۵، ۷(۱۳) https://www.mtlj.usc.ac.ir

حقوق فناوری‌های نوین

http://doi.org/10.22133/mtlj.2025.489698.1403

هوشمندسازی فرایند تفهیم اتهام با استفاده از شبکهٔ عصبی NLP با تأکید بر تشخیص عنوان مجرمانه

سعید گوهری

استادیار گروه حقوق، دانشکدهٔ ادبیات و علوم انسانی، دانشگاه زابل، زابل، ایران.

اطلاعات مقاله

مقالهٔ پژوهشی

تاریخ دریافت:
۱۴۰۳/۰۹/۲۰

تاریخ پذیرش:
۱۴۰۳/۱۲/۰۳

واژگان کلیدی:
تفهیم اتهام
شبکهٔ عصبی
پردازش زبان طبیعی
تشخیص عنوان مجرمانه
هوش مصنوعی

چکیده

فرایند تفهیم اتهام یکی از مراحل حیاتی در دادرسی کیفری است که نقشی اساسی در تضمین حقوق متهم ایفا می‌کند. باوجوداین، عواملی همچون حجم زیاد اطلاعات، پیچیدگی قوانین و استفاده از اصطلاحات فنی و تخصصی، تفهیم دقیق و سریع اتهام را به چالشی جدی تبدیل کرده است. این مقاله با بهره‌گیری از روش تحقیق ترکیبی، شامل تحلیل محتوا و پیاده‌سازی مدل‌های یادگیری عمیق مبتنی بر شبکه‌های عصبی پردازش زبان طبیعی (NLP)، به ارائهٔ راهکارهایی برای هوشمندسازی این فرایند می‌پردازد. تمرکز اصلی پژوهش بر تشخیص عنوان مجرمانه از داده‌های حقوقی در پرونده‌های قضایی است. یافته‌های پژوهش نشان می‌دهد که این فناوری با دقتی مقبول قادر به شناسایی و تفکیک دقیق عناوین مجرمانه است. بااین‌حال، چالش‌هایی همچون نیاز به داده‌های حجیم و دقیق برای آموزش الگوها، محدودیت در تطبیق الگوریتم‌ها با پیچیدگی‌های زبان حقوقی، و نگرانی‌های مربوط به حفظ حریم خصوصی و امنیت اطلاعات نیز وجود دارد. این دستاورد نوآورانه، با وجود چالش‌های ذکر شده، نه‌تنها موجب ارتقای دقت و سرعت در تفهیم اتهام می‌شود، بلکه به کاهش خطاهای انسانی و بهبود کلی عدالت کیفری کمک شایانی می‌کند.

***نویسندهٔ مسئول**
رایانامه: sgohari@uoz.ac.ir

نحوهٔ استناددهی:
گوهری، سعید (۱۴۰۵). هوشمندسازی فرایند تفهیم اتهام با استفاده از شبکهٔ عصبی NLP با تأکید بر تشخیص عنوان مجرمانه. حقوق فناوری‌های نوین، ۷(۱۳)، ۱۲۷-۱۴۸.

ناشر: دانشگاه علم و فرهنگ https://www.usc.ac.ir
شاپای الکترونیکی: ۳۸۳۶-۲۷۸۳

مقدمه

تفهیم اتهام یکی از مهم‌ترین مراحل دادرسی کیفری است که بر پایهٔ آن، مقام قضایی وظیفه اتهام منتسب به متهم را با شفافیت، دقت و در چهارچوب موازین قانونی به وی اعلام می‌کند. این فرایند به‌عنوان یکی از الزامات عدالت کیفری، از دو جنبه حائز اهمیت است: نخست، از منظر تضمین رعایت حقوق دفاعی که باعث می‌شود متهم از ماهیت اتهام و مستندات قانونی برای دفاع مؤثرمرتبط آگاهی کامل داشته باشد؛ همچنین این مهم به حفظ شفافیت و مشروعیت نظام قضایی می‌انجامد؛ چراکه تفهیم صحیح و دقیق اتهام از بروز سوءتفاهم‌های حقوقی، تفسیرهای نادرست از قانون و نقض حقوق متهم جلوگیری می‌کند. بااین‌حال، چالش‌هایی نظیر پیچیدگی زبان حقوقی، تنوع مواد قانونی و عناوین مجرمانه، گاه روند تفهیم اتهام را با ابهام مواجه می‌سازد و می‌تواند زمینه‌ساز صدور احکام ناعادلانه شود. ازاین‌رو، ارتقای دقت و استانداردسازی این فرایند با بهره‌گیری از ابزارهای نوین حقوقی و فناوری‌های پیشرفته، می‌تواند گامی مؤثر در راستای تضمین عدالت کیفری و صیانت از حقوق متهمان تلقی شود. بی‌تردید، استفاده از فناوری‌های نوین نقشی اساسی در تضمین عدالت کیفری به ارمغان می‌آرود. باوجوداین ، این فرایند با چالش‌های متعددی برای احراز صحیح عنوان مجرمانه به‌دلیل پیچیدگی‌های متون قانونی، شباهت‌های موجود میان برخی جرائم و تفسیرپذیری قوانین روبه‌روست که نه‌تنها ممکن است به صدور احکام غیرمنصفانه منجر می‌شود، بلکه حقوق دفاعی متهم را نیز به خطر می‌اندازد. برای مثال، شعبهٔ ۳۸ دیوان عالی کشور برابر با دادنامهٔ شماره ۹۵۰۹۹۷۲۸۱۱۶۰۰۴۲۸ مورخ ۱۳۹۵/۰۴/۳۱ رأی شمارهٔ ٤٢٨-۳۱/٤/۹۵ شعبهٔ اول دادگاه تجدیدنظر استان قزوین را، که براساس دادنامهٔ شمارهٔ ۱٦۰۷ صادره از شعبهٔ دادگاه بخش الف. آقای م.س. را نسبت به اتهام اخلال در نظم عمومی، به دو فقره حبس ۱۸ماهه و ۷٤ ضربه شلاق و بابت استفاده از دو فقره سلاح غیرمجاز به هفت ماه حبس تعزیری محکومیت یافته است با این استدلال که عمل ارتکابی متقاضی اعادهٔ دادرسی، با عنایت به مفاد دادنامهٔ بدوی تحت عنوان شرکت یا معاونت در نزاع دسته‌جمعی قابلیت تعقیب و مجازات داشته و صدور حکم به محکومیت مشارٌالیه به اتهام اخلال در نظم عمومی، وجاهت قانونی نداشته و فراقانونی تشخیص داده می‌شود، مردود اعلام کرده است.[1]

این چالش‌ها لزوم بازنگری در رویکردهای سنتی و استفاده از ابزارهای نوین، به‌ویژه فناوری‌های هوشمند نظیر پردازش زبان طبیعی (NLP)، را آشکار می‌سازد تا ازطریق استانداردسازی و افزایش دقت در فرایند تفهیم اتهام، عدالت کیفری به‌نحو شایسته‌تری محقق شود. اهمیت این موضوع ازآن‌روست که اشتباهات قضایی در تشخیص صحیح عنوان مجرمانه، یکی از موانع اصلی در تضمین حقوق قضایی متهمان در دادرسی منصفانه به‌شمار می‌رود. از همین رو، برخی از صاحب‌نظران تأکید دارند که تضمین حقوق افراد از سوی دولت‌ها نباید صرفاً به شناسایی و تصویب این حقوق در قوانین محدود شود؛ بلکه باید در عمل و در رویهٔ قضایی نیز رعایت شود تا ضمانت اجرایی مؤثری برای حفظ این حقوق فراهم شود (Mahmoudi Janaki & Jalil Piran, 2021, p. 314). در همین راستا، هرگونه خطای انسانی در تشخیص عنوان مجرمانه علاوه بر طولانی‌تر کردن فرایند دادرسی کیفری، به‌طور جدی حقوق متهم را تحت‌تأثیر قرار می‌دهد و به عواقب منفی و آسیب‌های جدی برای متهمان و خانواده‌های آنان منجر می‌شود (Brooks & Greenberg, 2021, p. 44) که این نقصان در تشخیص صحیح عنوان مجرمانه، اعتماد عمومی جامعه را نسبت به عملکرد نظام قضایی تضعیف می‌کند. بررسی‌ها نشان می‌دهد اکثر نظام‌های قضایی برای تشخیص عنوان مجرمانه تلاش می‌کنند تا با اتکا بر تحلیل شواهد و مدارک پرونده توسط مقامات قضایی مانند دادستان‌ها و قضات کیفری، ابتدا به جمع‌آوری دلایل و مستندات اثباتی و انتسابی شامل اظهارات شاهدان، مدارک مکتوب و شواهد جمع‌آوری‌شده از صحنهٔ جرم با هدف احصای رفتارهای ارتکابی می‌پردازند و پس از تحلیل‌های دقیق و فنی، عنوان اتهامی را مطابق با قوانین کیفری شناسایی می‌کنند. در این روش، به‌دلیل وجود خطای انسانی، امکان اشتباه در تشخیص عنوان جرم وجود دارد. از همین رو، حق تجدیدنظر علیه تصمیمات قضایی در قانون پیش‌بینی شده است (Tan, 2023, p. 89). باوجوداین، ظهور فناوری‌های نوین مانند هوش مصنوعی و پردازش زبان طبیعی (NLP) توانسته‌اند با تحلیل دقیق‌تر متون قانونی و

---

۱. برگرفته شده از:https://ara.jri.ac.ir/Judge/Text/34692

شناسایی رفتارهای ارتکابی براساس مستندات پروندهٔ قضایی، در زمینهٔ تحقق عدالت کیفری کمک شایانی کنند. برای نمونه، استفاده از فناوری‌های نوین مانند پردازش زبان طبیعی (NLP) در سیستم‌های قضایی به طور فزاینده‌ای باعث بهبود دقت، شفافیت و کارایی فرایندهای قانونی شده است. این فناوری، به ویژه در مراحل مختلف دادرسی کیفری، از جمله فرایند تفهیم اتهام، کاربردهای مهمی دارد که می‌تواند به نفع نظام‌های قضایی و حقوقی عمل کند. برای مثال، قوهٔ قضائیهٔ کشور کنیا از فناوری‌های پردازش زبان طبیعی (NLP) و هوش مصنوعی (AI) برای تجزیه‌وتحلیل اسناد حقوقی و پیش‌بینی نتایج پرونده‌ها استفاده می‌کند. این فناوری‌ها با دقت بالا، به خودکارسازی فرایندهای حقوقی و بهبود کارایی سیستم قضایی کمک می‌کنند (Obanda, 2022, p. 3). افزون‌براین، سیستم قضایی کشور هلند با استفاده از هوش مصنوعی، پیشنهادهای قضایی به دادگاه‌ها ارائه می‌دهد (Lim, 2021, p. 282). همچنین، اتحادیهٔ اروپا حمایت‌های گسترده‌ای در راستای به‌کارگیری از شبکهٔ NLP و فناوری‌های نوین مبتنی بر وب با هدف استخراج دانش از متون قانون برای رسیدگی به پرونده‌های قضایی صورت داده است (Robaldo et al., 2019, p. 114). به نظر می‌رسد، استفاده از فناوری NLP به‌دلیل قابلیت‌های پیشرفته در تجزیه‌وتحلیل متون حقوقی و استخراج اطلاعات ارزشمند، امکان استانداردسازی فرایند تفهیم اتهام، کاهش تبعیض ناشی از تفسیرهای متفاوت قوانین و شفافیت بیشتر در روند دادرسی را فراهم می‌سازد؛ به‌ویژه آنکه این فناوری مسیر دستیابی به شناسایی عنوان مجرمانهٔ صحیح را منطبق بر مدارک و مستندات پروندهٔ متهم در چارچوب قوانین کیفری برای آغاز فرایند تفهیم اتهام با کمترین خطاهای انسانی به ارمغان می‌آورد.

ازاین‌رو، هوشمندسازی فرایند تفهیم اتهام با استفاده از شبکه‌های عصبی NLP ضرورتی اجتناب‌ناپذیر در راستای تحقق عدالت قضایی به‌شمار می‌رود؛ زیرا بهره‌گیری از پردازش زبان طبیعی (NLP) می‌تواند راه را برای تشخیص عنوان مجرمانهٔ صحیح فراهم سازد. این فرایند به مقام قضایی کمک می‌کند تا عنوان اتهامی را به‌صورت صریح و روشن به متهم اعلام کند و او را از عملی که به خاطر آن سزاوار سرزنش و تعقیب کیفری است، آگاه سازد (Piran, 2021, p. 315). بنابراین، پژوهش حاضر نیز با هدف بررسی نقش این فناوری در افزایش دقت تشخیص عنوان مجرمانه و کاهش خطاهای قضایی نگارش شده و به دنبال پاسخ به این پرسش است که چگونه می‌توان از شبکهٔ عصبی NLP برای ارتقای فرایند تشخیص عنوان مجرمانه در تفهیم اتهام بهره گرفت؟

در راستای پاسخ به پرسش حاضر، در گام نخست، پیشینهٔ تحقیق به بررسی مطالعات مرتبط با هوش مصنوعی در نظام‌های قضایی و تحلیل کارکردهای NLP برای تشخیص عنوان مجرمانه می‌پردازد تا ضرورت استفاده از این فناوری را در نظام قضایی ایران آشکار سازد. سپس، فرایند پیش‌پردازش داده‌های حقوقی شامل مراحل آماده‌سازی و پالایش اطلاعات برای استفاده در مدل‌های یادگیری عمیق تبیین شده است. در گام بعدی، تحقیق بر تشخیص عنوان اتهامی با بهره‌گیری از معماری شبکهٔ عصبی چندلایه متمرکز است که ساختار الگو و نحوهٔ استخراج اطلاعات از متون حقوقی را ارائه می‌دهد و پس از آن، مراحل آموزش شبکهٔ عصبی برای تشخیص عناوین مجرمانه و تحلیل عملکرد آن براساس مجموعه‌داده‌ها تشریح می‌شود. در ادامه، چگونگی اعتبارسنجی عناوین اتهامی شناسایی شده و دقت عملکرد سیستم ارزیابی شده است. درنهایت، به چالش‌های قانونی، ساختاری و اجرایی این فناوری در نظام قضایی ایران پرداخته شده است. این ساختار با هدف پوشش همه‌جانبهٔ ابعاد نظری و عملی پژوهش تدوین شده و سعی بر آن بوده است تا نقش شبکه عصبی NLP در ارتقای فرایند تفهیم اتهام به‌طور جامع تبیین شود.

۱. پیشینه تحقیق

ژونگ[1] و همکاران (2020) در تحقیقی با عنوان «چگونه NLP از سیستم حقوقی سود می‌برد: خلاصه‌ای از هوش مصنوعی قانونی» بر اهمیت استفاده از فناوری هوش مصنوعی در انجام هرچه بهتر وظایف قانونی دستگاه قضایی تأکید کردند. یافته‌های پژوهش آنان بر این مهم تأکید دارد که استفاده از شبکهٔ عصبی NLP تخصصان حقوق را پیچ‌وخم کاغذبازی در رسیدگی‌های قضایی رها می‌بخشد. این محققان ضمن اشاره به

1. Zhong

تاریخچه، وضعیت فعلی و جهت‌گیری‌های آیندهٔ استفاده از فناوری هوش مصنوعی در حوزهٔ حقوق، نتایجی را از تجزیه‌وتحلیل چند آزمایش دراین‌خصوص ارائه داده‌اند که نشان می‌دهد استفاده از این فناوری به عدالت کیفری کمک شایانی می‌کند.

لومبو[1] و همکاران (2022) در مقاله‌ای با عنوان «تشخیص و تجزیه‌وتحلیل جرم از پیام‌های رسانه‌های اجتماعی با استفاده از تکنیک یادگیری ماشینی و پردازش زبان طبیعی» به این موضوع پرداختند که رسانه‌های اجتماعی تأثیر شایان توجهی در ماهیت و میزان جرم و جنایت در جوامع داشته‌اند. یافته‌های پژوهش آنان نشان داد که ناشناس ماندن کاربران و نبود منابع کافی برای مقابله با جرایم سایبری از عوامل اصلی افزایش این فعالیت‌ها، به‌ویژه در آفریقای جنوبی، هستند. این گروه با استفاده از الگوریتم‌های پردازش NLP برای پردازش متون و مدل‌های یادگیری ماشینی نظیر ماشین‌های بردار پشتیبان و طبقه‌بندهای جنگل تصادفی برای طبقه‌بندی داده‌ها نشان دادند. گفتنی است دقت ماشین بردار پشتیبان در شناسایی جرایم متنی، ۸۶ درصد و دقت جنگل تصادفی، ۷۲ درصد بوده است.

اوسوریو و بلترن[2] (2020) در تحقیقی با عنوان «افزایش تشخیص سازمان‌های جنایت‌کار در مکزیک با استفاده از NLP و ML» به دنبال ایجاد یک پایگاه‌دادهٔ جغرافیایی برای یادگیری ماشین (ML) و پردازش زبان طبیعی (NLP) با هدف شناسایی حضور سازمان‌های جنایی مکزیکی بوده‌اند. یافته‌های مطالعهٔ آن‌ها نشان داد که این دو شبکه با کمک داده‌های باکیفیت بالا درمورد گروه‌های جنایت‌کار برای اطلاع‌رسانی به تحلیل علمی و سیاستی در زمینهٔ خشونت شدید، مانند مکزیک پاسخ می‌دهد. این برنامهٔ کاربردی علوم اجتماعی محاسباتی با پشتیبانی از ابزارهای ML و NLP، مجموعهٔ گسترده‌ای از اخبار نوشته شده به زبان اسپانیایی را برای ردیابی حضور خشونت‌آمیز MCOها پردازش می‌کند.

بلوچ‌فرد (2022) در مقاله‌ای با عنوان «پردازش زبان طبیعی و تحلیل داده‌های حقوقی در پرونده‌ها و گزارش پزشکی قانونی با استفاده از هوش مصنوعی»، با بررسی اثر پردازش متون حقوقی، گزارش‌ها و شرح پروندهٔ جنایی با استفاده از «پردازشگر زبان طبیعی»، به این نتیجه دست یافت که با استفاده از هوش مصنوعی می‌توان یک نظام مؤثر از تلفیق بین حقوق، فناوری و پزشکی برای بهبود روند کار در سازمان پزشکی قانونی ارائه کرد. رهیافت‌های حاصل از این پژوهش نشان داد که در برخی از کشورها، تلاش‌های گسترده‌ای در زمینهٔ توسعهٔ این فناوری برای تجزیه‌وتحلیل داده‌های حقوقی صورت گرفته است.

۲. پیش‌پردازش اولیهٔ داده‌های حقوقی

نخستین گام در فرایند هوشمندسازی تفهیم اتهام، آماده‌سازی داده‌های متنی نظیر متون قانونی و اوراق قضایی در قالبی است که امکان درک و تجزیه‌وتحلیل برای شبکهٔ پردازش زبان را فراهم سازد. این فرایند باعث می‌شود تحلیل‌های پیچیده در کوتاه‌ترین زمان (Mumcuoğlu et al., 2021, p. 2) ازطریق شبکه امکان‌پذیر شود؛ اما تحقق این هدف، نیازمند انجام مجموعه‌ای از اقدامات اولیه است که از آن به‌عنوان پردازش اولیهٔ داده‌ها برای ماشین‌های یادگیری یاد می‌شود. در واقع، پیش‌پردازش متن، فرایند به دست آوردن ارزش معنادار از مجموعه دادهٔ متن است (Yang, 2022, p. 2). توضیح آن‌که به‌دلیل وجود اشتباهات نگارشی و تایپی در متون قضایی، ماشین‌های یادگیری هوش مصنوعی تحت‌تأثیر کیفیت پایین داده‌ها نمی‌توانند عملکرد قابل دفاعی داشته باشند. درنتیجه، بی‌توجهی به سازوکار مناسب در این مرحله باعث می‌شود شبکهٔ عصبی (NLP) در تشخیص عنوان مجرمانه با خطا مواجه شود؛ بنابراین پیش‌پردازش اولیهٔ داده‌های قضایی، نوعی ریل‌گذاری برای هدایت داده‌های باکیفیت در مسیر لایهٔ ورودی به شبکه به‌شمار می‌رود. در ادامه، سعی بر آن است تا گام‌های کلیدی پیش‌پردازش داده‌های حقوقی برای تشخیص عنوان مجرمانه تبیین شود.

1. Lombo
2. Osorio & Beltran

۱-۲. پاک‌سازی و نرمال‌سازی[1] داده‌های حقوقی

همان‌طور که در سطور بالا اشاره شد، موفقیت شبکه در تشخیص صحیح عنوان اتهامی به‌شدت متأثر از کیفیت تحلیل داده‌های قضایی در لایهٔ ورودی این شبکه است؛ زیرا مفاهیم و معانی موجود در سند (اوراق پرونده) ازطریق تحلیل اطلاعات و برقراری ارتباط میان آن‌ها قابل‌استخراج است (Jalali Shirjani & Shirzad, 2021, p. 161). بااین‌حال، به‌دلیل نقش‌آفرینی اشخاص متعدد در شکل‌گیری و رسیدگی به پرونده‌های کیفری نظیر مقامات قضایی، وکلای دادگستری، شاکی، مظنون، ضابطین، کارشناسان دادگستری، شهود و غیره و همچنین، به‌دلیل واژگان فراوان و ویژگی‌های منحصربه‌فرد زبان فارسی، وجود اشتباهات تایپی، علائم نگارشی اضافی یا سایر ناهماهنگی‌های موجود در ساختار ادبی متون حقوقی دور از انتظار نیست. برای مثال، ممکن است در متن‌های مربوط به اوراق پروندهٔ قضایی نظیر صورت‌جلسهٔ استماع اظهارات شهود، به‌جای واژهٔ «سرقت» به اشتباه « سر اقت» نوشته شده باشد یا این‌که در لایحهٔ دفاعیهٔ تقدیمی از سوی متهم، علائم نگارشی اضافه وجود داشته باشد؛ برای نمونه، به‌جای علامت نگارشی «؛» از علامت نگارشی«؛،» استفاده شده باشد. این معضل، چالش‌های اساسی را برای تصحیح خطای کلمات واقعی ایجاد می‌کند (Hickman et al., 2022, p. 221)؛ زیرا بهبود دقت مدل، کاهش پیچیدگی محاسبات داده‌ها و افزایش قابلیت تفسیر نتایج، ارتباط مستقیم باکیفیت داده‌های ورودی به شبکه دارد و می‌تواند بر قدرت آماری تحلیل‌های بعدی و اعتبار بینش‌های حاصل از متن‌کاوی نیز اثرگذار باشد (Hickman et al., 2022, p. 114). اما پرسش اساسی این است که تصحیح خطاهای نگارشی چگونه امکان‌پذیر است؟ به نظر می‌رسد این چالش با دو روش اصلاح شود: در روش نخست، می‌توان از مدل‌های مبتنی بر توالی‌به‌توالی[2]، که تکنیکی رایج در NLP است، بهره گرفت. این روش براساس شبکه‌های عصبی در زمینهٔ فعالیت‌هایی مانند ترجمهٔ ماشینی، خلاصه‌سازی متن و اصلاح نگارشی به کار می‌رود (Dongbo et al., 2023, p. 2) و می تواند موثر واقع شود.

اما در روش دوم، از مدل‌های مبتنی بر مدل‌های ترنسفورمر استفاده می‌شود. این مدل‌ها دوبخشی هستند و می‌توانند با شناسایی کلمات و عبارات نادرست به درک معنای جملات برای تشخیص عنوان اتهامی کمک شایانی کنند. برای مثال در جرم توهین مضوع مادهٔ ۶۰۸ قانون مجازات اسلامی مصوب ۱۳۹۲ اگر در متن پرونده قید شده باشد که «متهم در یک جلسهٔ عمومی به شاکی الفاظ توهین‌آمیز نسبت داده و باعث هتک حرمت وی شده است و این الفاظ شامل عباراتی نظیر دروغگو، خائن و بی‌شرم باشند» فرایند نرمال‌سازی عباراتی مانند «الفاظ توهین‌آمیز»، «هتک حرمت»، «دروغگو»، «خائن» و «بی‌شرم» به مدل کمک می‌کند تا جرم «توهین» را تشخیص دهد. این اصطلاحات ممکن است به اشکال مختلف در متون ظاهر شوند؛ اما با استفاده از نرمال‌سازی، تمام این موارد به یک عنوان مجرمانهٔ استاندارد مرتبط می‌شوند. مثال دیگری که در این خصوص می‌توان به آن اشاره کرد، جرم حمل مواد مخدر موضوع قانون مبارزه با مواد مخدر مصوب ۱۳۶۷ است. اگر خلاصهٔ متن پرونده نشان دهد که «متهم هنگام عبور از مرز، مقدار قابل توجهی از مواد مخدر شامل شیشه و هروئین را در خودرو خود مخفی کرده بود». فرایند نرمال‌سازی عباراتی مانند: «حمل مواد مخدر»، «مخفی کردن»، «شیشه»، و «هروئین» به مدل امکان می‌دهند که رفتارهای ارتکابی را تحت عنوان «حمل مواد مخدر» دسته‌بندی کند. ازهمین‌رو، تشخیص عنوان مجرمانه با استفاده از شبکه‌های پردازش زبان طبیعی (NLP) نیازمند یکسان‌سازی و استانداردسازی داده‌ها برای تحلیل و پردازش است. این مرحله، پلی برای استخراج واژگان کلیدی در حوزهٔ استخراج متن، بازیابی اطلاعات و پردازش طبیعی زبان (Jafari Powersi et al., 2019, p. 351) به‌شمار می‌رود. این فرایند، علاوه‌بر این‌که به یک‌دست کردن متون حقوقی برای پردازش کمک می‌کند، باعث می‌شود متون حقوقی، که دارای ابهامات و کلمات چندمعنایی هستند، با بهره‌گیری از تکنیک‌های پایه‌سازی به متونی قابل‌فهم برای شبکه تبدیل شوند. برای مثال، نرمال‌سازی اصطلاحات قانونی مرتبط با جرمی نظیر ضرب‌وجرح عمدی به این صورت است که شبکهٔ «ضرب‌وشتم» و «کتک‌کاری» را تحت عنوان ضرب‌وشتم شناسایی می‌کند. همچنین برای پایه‌سازی کلمات برای این‌که به شکل ساده‌تری تبدیل شوند، می‌توان از این فرایند استفاده کرد. برای نمونه، کلمات موجود در یک پروندهٔ قضایی شامل «متهمان»،

1. Normalization
2. Seq-To- Seq

«متهمی»، «متهمین» برای شبکه به «متهم» تبدیل شود. این فرایند به شناسایی دقیق‌تر مفاهیم حقوقی و تحلیل الگوهای زبانی مشابه پیرامون توصیف عناوین مجرمانه کمک شایانی می‌کند و باعث بهبود ارتقای کارایی مدل‌های NLP می‌شود.

۲-۲. توکن‌سازی و تبدیل داده‌های حقوقی به بردارهای عددی

مدل‌های یادگیری ماشین برای این‌که قادر باشند متون حقوقی و قضایی را به داده‌های قابل درک معنایی تبدیل کنند، نیازمند آن هستند که واژگان و کلمات را به اجزایی کوچک‌تر و معنادار تبدیل کنند. در این فرایند، جملات یا اسناد مرتبط با پرونده‌های کیفری پیش از ورود به مدل‌های محاسباتی باید به بخش‌های کوچک‌تری شکسته شوند. این تقسیم‌بندی کمک می‌کند اصطلاحات پیچیده و عبارات تخصصی به واحدهایی ساده‌تر و قابل‌فهم‌تر برای مدل تبدیل شوند (Michelbacher, 2013, p. 13). برای نمونه، جمله‌ای مانند «متهم یک قطعه‌سنگ یاقوت برمه به وزن ۸۲/۸۱ گرم از گاوصندوق دزدی کرده است» می‌تواند به توکن‌های زیر تقسیم شود: («متهم»، «یک»، «قطعه»، «سنگ»، «یاقوت»، «برمه»، «به»، «وزن»، «۸۲»، «/»، «۸۱»، «گرم»، «از»، «گاوصندوق»، «دزدی»، «کرده»، «است»). این توکن‌ها در مراحل بعدی، مانند استخراج ویژگی‌ها و مدل‌سازی، به کار گرفته می‌شوند و به مدل کمک می‌کنند تا روابط معنایی و الگوهای پنهان در داده‌ها را شناسایی کرده و عنوان مجرمانهٔ مربوط به متن را با دقت بیشتری تشخیص دهد.

البته باید توجه داشت که دقت در توکن‌سازی امری ضروری است؛ زیرا هرگونه خطا در این مرحله ممکن است به نتایج نادرست منجر شود (Wang et al., 2024, p. 3). یکی از راه‌های بهبود این فرایند، تمرکز مدل بر کلمات و اصطلاحات کلیدی است. برای مثال، در متن‌های حقوقی، برخی کلمات مانند «از» یا «با» معمولاً اهمیت چندانی ندارند و می‌توان آن‌ها را از متن حذف کرد. در مقابل، واژه‌ها و اصطلاحات خاص حقوقی مانند «دزدی»، «جرم»، «قانون» و «محکوم» از اهمیت بیشتری برخوردارند و می‌توانند به منزلهٔ ویژگی‌های کلیدی برای شناسایی عنوان مجرمانه استفاده شوند. این تمرکز بر واژه‌های کلیدی به مدل‌های NLP اجازه می‌دهد تحلیل دقیق‌تری از متن ارائه دهند و روابط معنایی را بهتر شناسایی کنند؛ چراکه در مدل‌های یادگیری پردازش زبان طبیعی، کلمات به‌صورت واحدهای مستقل در نظر گرفته می‌شوند و هیچ شباهت ذاتی میان آن‌ها وجود ندارد. به همین دلیل، برای درک روابط معنایی میان واژه‌ها (Pakzad & Anna Louis, 2022, p. 2)، ابتدا حجم گسترده‌ای از داده‌ها فشرده می‌شود (Mikolov et al., 2013, p. 2) و سپس این داده‌ها به بردارهای عددی تبدیل می‌شوند. برای مثال، ممکن است مدل یادگیری واژهٔ «خیانت در امانت» را به بردار عددی (۰٫۱۲, ۰٫۷۵, ۰٫۳۳...) و واژهٔ «کلاهبرداری» را به بردار عددی (۰٫۸۵, ۰٫۲۵, ۰٫۷۰...) تبدیل کند. این بردارها به شبکهٔ عصبی کمک می‌کنند تا روابط معنایی میان واژه‌ها را درک کند و عباراتی را که بیشترین شباهت به عنوان مجرمانه دارند، برای رفتار ارتکابی شناسایی و پیشنهاد دهد. ازهمین‌رو، تبدیل واژه‌ها به بردارهای عددی، امکان تحلیل دقیق‌تر و مؤثرتر متن را فراهم می‌کند و نقش مهمی در بهبود فرایند شناسایی عناوین مجرمانه و درک بهتر از داده‌های حقوقی دارد.

۲-۳. پدگذاری توالی[1] و مقیاس‌سازی[2] بردارهای عددی

شبکه‌های عصبی (NLP) برای این‌که بتواند داده‌های عددی را به شکل ماتریس‌های قابل‌محاسبه مورد پردازش قرار دهند، نیازمند فرایندی هستند که بر مبنای آن توالی‌های عددی جهت ورود داده‌ها به مدل‌های یادگیری عمیق به یک طول ثابت دست یابند. به‌عبارت‌دیگر، تمامی داده ها جهت ورود به شبکه باید دارای طول یکسانی باشند. به‌عنوان نمونه، در یک پرونده کیفری، اظهارات شهود ممکن است به‌صورت جملاتی با طول‌های متفاوتی بیان شوند. برای مثال: «شاهد اول گواهی می‌دهد که: «مدیر مالی مبلغی معادل ۵۰۰ میلیون تومان را از حساب شرکت برداشت کرد و شاهد دوم گواهی می‌دهد که: «مدیر مالی با جعل اسناد مالی، حق شرکت را تضییع کرد». همان‌طور که ملاحضه می‌شود، طول این دو جمله یکسان نیست. به‌عبارت‌دیگر، جملهٔ اول شامل یازده کلمه و جملهٔ دوم شامل نه کلمه است. در این شرایط، مدل یادگیری عمیق برای این‌که قادر

1. Sequence Padding
2. Data Scaling

باشد داده‌ها را به‌صورت استاندارد پردازش نماید باید طول داده ها را برای ماشین یادگیری به توالی‌های یکسان تبدیل کند. این فرایند، زمینهٔ محاسبهٔ صحیح داده‌ها را از رهگذر ماتریس‌های قابل محاسبه فراهم می‌سازد، به نحوی که بسترهای لازم جهت توصیف جمله توسط ماتریس فراهم می‌شود (Giménez, 2019, p. 316). راه‌حلی که برای برون‌رفت از این وضعیت ارائه شده است فرایند پدگذاری است که براساس آن، در مواردی که طول دو جملهٔ به‌هم‌پیوسته کوتاه‌تر از حداکثر طول دنباله باشد، به انتهای جملات کوتاه‌تر بالشتک اضافه می‌شود تا جاهای خالی پر شود (Zhang et al., 2021, p. 92). برای نمونه، اگر در پرونده‌ای مدیر مالی یک شرکت به استناد مادهٔ ٦٧٤ قانون مجازات اسلامی مصوب ١٣٩٢ متهم به خیانت در امانت شود. اگر جملات شهود، بدین شکل باشند: شاهد اول: «مدیر مالی مبلغ ٧٠٠ میلیون تومان از حساب شرکت برداشت کرد و به حساب شخصی خود انتقال داد.» شاهد دوم: «مدیر مالی اسناد حسابداری شرکت را جعل کرد.» در اینجا، جملهٔ اول هفده کلمه و جمله دوم هشت کلمه دارد. اگر طول بزرگترین جمله متشکل ازعدد باشد، به جمله اول یک «پد» و به جمله دوم، سه «پد» (که معمولاً با صفر جایگزین می شوند) مانند جدول زیر اضافه می‌شود تا هر دو ورودی از طول یکسانی برخوردار شوند.

جدول ١ : پدگذاری جملات متن‌های قضایی

ردیف	جملهٔ قبل از پدگذاری	جملهٔ بعد از پدگذاری
١	مدیر مالی مبلغ ٧٠٠میلیون تومان از حساب شرکت برداشت کرد و به حساب شخصی خود انتقال داد.	جمله اول: {مدیر، مالی، مبلغ، ٧٠٠، میلیون، تومان، از، حساب، شرکت، برداشت، کرد، و، به، حساب، شخصی، خود، انتقال، داد}
٢	مدیر مالی اسناد حسابداری شرکت را جعل کرد.	{مدیر، مالی، اسناد، حسابداری، شرکت، را، جعل، کرد، ٠، ٠، ٠، ٠}

بااین‌حال، اگرچه پدگذاری فرایندی مؤثر برای یکنواخت‌سازی طول جملات ورودی در شبکه‌های عصبی است؛ ممکن است باعث ایجاد خطاهایی در تشخیص عنوان مجرمانه شود. دلیل این امر آن است که مدل‌های یادگیری عمیق در پردازش داده‌های متنی ممکن است به‌جای تحلیل معنایی، بر مقادیر آماری مانند تعداد تکرار کلمات تمرکز کنند. برای مثال، در پرونده‌ای مرتبط با خیانت در امانت موضوع مادهٔ ٦٧٤ قانون مجازات اسلامی مصوب ١٣٩٢، ممکن است دو شاهد بدین شکل شهادت دهند: شاهد اول: «مدیر مالی سه بار خیانت در امانت انجام داد و دو بار اختلاس کرد.» و شاهد دوم: «مدیر مالی درمجموع پنج بار خیانت در امانت انجام داد و تنها یک بار اختلاس کرد.» در این حالت، شبکه ممکن است براساس تعداد تکرار کلمات، جرم خیانت در امانت را مهم‌تر تلقی کند و جرم اختلاس را نادیده بگیرد؛ حتی اگر ماهیت پرونده بر اختلاس متمرکز باشد. این چالش زمانی جدی‌تر می‌شود که داده‌ها از لایه‌های عمیق شبکه عبور کرده و ویژگی‌های آماری بر تحلیل معنایی غالب شوند (Sharma, 2022, p. 31). این مشکل در شبکه‌های عمیق، که داده‌ها از لایه‌های متعدد عبور می‌کنند و از توابعی مانند «Sigmoid»بهره می‌برند، تشدید می‌شود؛ زیرا مقادیر ورودی بزرگ می‌توانند تابع را به‌سمت اشباع سوق دهند و باعث توقف یادگیری شوند. برای رفع این مشکل، مقیاس‌بندی داده‌ها ضروری است. این فرایند شامل نرمال‌سازی داده‌ها در بازه [٠,١] است که به مدل کمک می‌کند تا وابستگی‌های معنایی را بهتر درک کند و به تکرارهای بی‌مورد وزن کمتری بدهد. برای مثال، فرض کنید در پرونده‌هایی با ماهیت جرائمی نظیر، جعل و کلاهبرداری موضوع مواد ٥٢٣ قانون مجازات اسلامی و مادهٔ ١ قانون تشدید مجازات مرتکبان اختلاس، ارتشا و کلاهبرداری مصوب ١٣٦٧، مبالغ متفاوتی توسط متهم از حساب یک شرکت برداشته شده است: «جعل: ٥٠ میلیون تومان وکلاهبرداری: ١ میلیارد تومان.» در چنین حالتی، اگر در مرحلهٔ پردازش اولیهٔ داده‌ها، توجه لازم به مقیاس‌بندی صورت نگیرد، مدل ممکن است جرم جعل را به این دلیل نادیده بگیرد که مبلغ عددی آن در مقایس با مبلغ عددی جرم کلاهبرداری کمتر است؛ ازاین‌رو با مقیاس‌بندی پیش‌پردازش داده‌ها می‌توان در عملکرد الگوریتم‌های یادگیری ماشین تأثیر مثبت ایجاد کرد (Ambarwari et al., 2020, p. 117). این فرایند در افزایش دقت و پایداری شبکه‌های عصبی در تشخیص عناوین مجرمانه نقشی اصلی ایفا می‌کند. این روش با تعدیل داده‌ها و جلوگیری از اشباع توابع شبکه، مدل را به‌سمت یادگیری بهتر و تحلیل دقیق‌تر هدایت می‌کند. درنتیجه، خروجی مدل پایدارتر، قابل‌اعتمادتر و نزدیک‌تر به واقعیت خواهد بود.

۳. تشخیص عنوان اتهامی با معماری شبکهٔ عصبی چندلایه

۳-۱. لایهٔ ورودی داده‌های حقوقی

برای این‌که داده‌های خام (متن‌های حقوقی و قضایی) به شکلی تبدیل شوند که شبکهٔ عصبی قادر به پردازش آن باشد، باید در طراحی لایهٔ ورودی شبکه دقت لازم و کافی صورت پذیرد؛ زیرا تحلیل مناسب داده‌ها و ارائهٔ نتایج دقیق در تشخیص عنوان مجرمانه منوط به دقت در طراحی لایهٔ ورودی خواهد بود. بااین‌حال، طراحی لایهٔ ورودی به‌منظور پردازش داده‌های مرتبط با متون حقوقی در مقایسه با اسناد حوزه‌های دیگر مرحله‌ای چالش‌برانگیز است (Chen et al., 2022, p.2). برای مثال، در یک پرونده کیفری با ماهیت فرار مالیاتی یا قاچاق مواد مخدر، اگر اسناد متنی، مانند صورت‌جلسهٔ کشف ضابطین قضایی، اظهارات شاهدان و اظهارات مظنون یا متهم به‌عنوان ورودی شبکه در نظر گرفته شود، به‌دلیل آن‌که هر یک از این اسناد شامل تعداد زیادی از کلمات هستند، باید سازوکاری پیش‌بینی شود تا اسناد یاد شده پس از تبدیل به قالب‌های عددی ازطریق نودهای لایهٔ ورودی وارد شبکهٔ عصبی شوند. بااین‌حال، با توجه به این‌که لایهٔ ورودی با حجم گسترده‌ای از داده‌های توکن‌شده روبه‌رو خواهد بود، لازم است سازوکاری در طراحی لایهٔ ورودی مدنظر قرار گیرد تا مدل از پردازش داده‌های غیرضروری خودداری کند؛ بنابراین در معماری ورودی شبکه به دو امر حیاتی باید توجه شود: نخست، توجه به فرایند ماسک‌گذاری است. ماسک‌گذاری باعث می‌شود که لایهٔ ورودی بر داده‌هایی تمرکز کند که قرار است شبکه آن‌ها را در محاسبات لحاظ کند.

موضوع دوم این است که در طراحی لایهٔ ورودی شبکه، ضروری است تا نودهای ورودی شبکه از استاندارد کیفی و کمی لازم برخوردار باشند؛ زیرا سلول‌های ورودی شبکه پیچیدگی بسیار کمی دارند و به همین منظور، فضای معماری شبکه با محدودیت مواجه است (Klyuchnikov et al., 2022, p. 45737). برای رفع این چالش، دو رویکرد متداول است: در رویکرد نخست، از مدل‌های مبتنی بر کلمات[1] استفاده می‌شود که در آن، تعداد نودها در لایهٔ ورودی براساس تعداد کلمات تعیین می‌شود و براساس مدل‌های تعبیه شدهٔ نمایش برداری، کلمات توزیع می‌شوند (Kumar Singh et al., 2020, p. 1)؛ اما در رویکرد دوم، از مدل‌های مبتنی بر ویژگی‌ها[2] در تعیین تعداد نودها بهره گرفته می‌شود و تعداد نودها برابر با تعداد ویژگی‌هایی استخراج شده از اوراق یک پرونده قضایی تعیین می‌شوند؛ زیرا استخراج کلمات کلیدی از متن با بهره‌گیری از الگوریتم‌های تحلیل مقایسه‌ای امکان‌پذیر خواهد بود (Yang, 2022, p. 4). برای مثال، اگر در یک پروندهٔ قضایی با ماهیت جرائم مالی کارکنان دولت پنجاه‌هزار لغت وجود داشته باشد، اما ۱۲۰۰ لغت در کل پرونده به‌عنوان یک ویژگی تشخیص عنوان مجرمانه جرم خاص در نظر گرفته شده باشد، تعداد نودها همان ۱۲۰۰ لغت برای ورودی شبکه خواهد بود. این فرایند لایهٔ ورودی را به‌شکلی طراحی می‌کند که با توجه به ماهیت داده‌های زبان طبیعی، مدل بتواند بهینه‌ترین استفاده را از اطلاعات معنایی و ساختاری متن داشته باشد.

۳-۲. لایه‌های پنهان شبکه[3]

لایه‌های پنهان شبکهٔ عصبی (NLP) نقشی بی‌بدیل در شناسایی الگوهای مجرمانه ایفا می‌کنند. این لایه‌ها معمولاً میان لایهٔ ورودی و لایهٔ خروجی قرار می‌گیرند و به‌عنوان یک شبکهٔ عصبی عمیق، در زیربخش یادگیری ماشین برای طبقه‌بندی داده‌ها استفاده می‌شوند (Abul Bashar, 2019, p.76). نقش این لایه‌ها برقراری روابط معنایی میان کلمات، متن‌ها و جملات است (Young, 2018, p. 56). در واقع، شبکه با استفاده از لایه‌های پنهان قادر خواهد بود تا روابط معنایی پیچیده میان اصطلاحات حقوقی، قوانین و شواهد ارائه شده در مراجع قضایی (دادسرا و محاکم کیفری) را برای تشخیص عنوان مجرمانه کشف کند. برای مثال، در جرم کلاهبرداری موضوع مادهٔ (۱) قانون تشدید مجازات مرتکبان ارتشا و اختلاس و کلاهبرداری «توسل توأم با سوءنیت به وسایل متقلبانه یا عملیات متقلبانه برای بردن مال دیگری»، یکی از شرایط اساسی برای احراز

1. Word Embeddings
2. Count Vectorization
3. Hidden Layer

جرم کلاهبرداری به‌شمار می‌رود. علاوه‌براین، ترکیب‌های معنایی کلماتی مثل «فریب خوردن»، «سوءنیت»، و «مستندات جعلی» از دیگر ویژگی‌های مربوط به این جرم است که ضرورت دارد به‌طور نظام‌مند و هوشمند ازطریق لایهٔ پنهان شبکه پردازش شوند تا از این رهگذر با برقراری روابط معنایی میان داده‌ها، تشخیص عنوان مجرمانه میسر شود؛ اما تحقق این هدف نیازمند توجه به دو نکتهٔ مهم است: اول این‌که تعداد نودهای لایه‌های پنهان باید به‌نحوی تنظیم شود که شبکه قادر به استخراج ویژگی‌های لازم از داده‌های ورودی باشد. برای مثال، در یک پروندهٔ کیفری با ماهیت «سوءاستفاده از سفیدمهر» (موضوع مادهٔ ۶۷۳ قانون مجازات اسلامی)، داده‌های ورودی به شبکه ممکن است شامل متن‌های طولانی از مطالب تعهدآور روی چندین سند مالی باشد که نیاز به تحلیل دقیق توسط شبکه وجود دارد. در این حالت، لایهٔ پنهان باید قادر به تجزیه‌وتحلیل عمیق مفاهیم باشد تا بتواند ارتباط میان مطالب تعهدآور و انطباق آن با الگوهای مجرمانهٔ مرتبط با این جرم را به سرانجام مقصود برساند.

همچنین یکی از مهم‌ترین موضوعاتی که در طراحی لایهٔ پنهان باید به آن دقت کافی شود، موضوع انتخاب تابع مناسب برای شناسایی الگوهای رفتاری مجرمانه است. برای مثال، اگر اطلاعات ورودی به شبکه شامل جرائمی نظیر، سرقت، کلاهبرداری، قتل، نزاع دسته‌جمعی، قاچاق کالا و مواردی از این قبیل باشد، این پرسش مطرح می‌شود که لایهٔ پنهان شبکه چگونه قادر به شناسایی الگوهای مجرمانه برای تشخیص عنوان مجرمانه خواهند بود؟ به نظر می‌رسد نودهای تعبیه شده در لایه‌های پنهان شبکه برای این‌که بتوانند روابط پیچیده میان داده‌ها را پردازش و شناسایی کنند باید بر ویژگی‌های داده‌های آموزش داده شده به شبکه در شناسایی عنوان مجرمانهٔ هر جرم تمرکز کنند. این مدل‌سازی‌ها در لایه‌های پنهان شبکه با استفاده از توابع خطی و غیرخطی صورت می‌پذیرد؛ زیرا توابع ابزار اصلی توصیف دنیای واقعی به زبان ریاضی هستند (Rafipour & Mohseni, 2019, p. 57) که با برقراری رابطهٔ ریاضی میان مجموعه‌ای از داده‌های ورودی به شبکه نقشی اساسی در شناسایی عنوان مجرمانه ایفا می‌کنند. باوجوداین، تشخیص الگوهای مجرمانه به‌دلیل پیچیدگی در روابط میان داده‌ها به‌آسانی امکان‌پذیر نیست؛ چراکه ویژگی‌های مرتبط با اوصاف مجرمانه در متن‌های قضایی به‌صورت پراکنده توزیع شده است و مدل‌های خطی به‌دلیل محدودیت‌هایی که دارند، قادر به شناسایی دقیق عناوین مجرمانه نخواهند بود.

ازاین‌رو، استفاده از تابع غیرخطی[1] به‌منظور برقرار کردن ارتباط منطقی میان مجموعه‌ای از اقدامات متهم و مدارک موجود در پروندهٔ قضایی با مواد قانون مفید به نظر می‌رسد؛ چراکه لایه‌های پنهان شبکه با کمک تابع غیرخطی ویژگی‌های داده‌های ورودی مربوط به هر جرم را ازطریق مقایسهٔ تفاوت‌ها و ویژگی‌ها دقیق میان داده‌های ورودی را با کمترین خطا تشخیص می‌دهند. بااین‌حال، ازآنجاکه داده‌های مربوط به الگوهای رفتاری هر جرم دارای ویژگی‌های فنی و تخصصی حقوقی هستند، شبکه با استفاده از یک مدل ریاضی مناسب برای ایجاد مرکزیت داده‌ها در اعداد بین (۱- و ۱) و با بهره‌گیری از تابع Tanh، براساس پایهٔ مدل ریاضی $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ قادر به تشخیص عنوان مجرمانه خواهد بود. برای نمونه، در یک پروندهٔ کیفری با ماهیت جرم اختلاس، برخی از جملات ممکن است شامل تفاوت‌های معنایی دقیق میان «خطای محاسبات در ارقام مالی» و «سوءنیت در استفاده از تنخواه دولتی» باشد که با استفاده از تابع Tanh، شبکه قادر به تشخیص تفاوت‌های موجود خواهد بود.

۳-۳. لایهٔ خروجی شبکه

لایهٔ خروجی شبکه نقش کلیدی در تشخیص عنوان مجرمانه دارد. ازاین‌رو، در طراحی آن باید دقت لازم صورت گیرد تا خروجی شبکه انعکاسی از عنوان مجرمانهٔ صحیح بر پایهٔ پردازش ویژگی‌های داده‌های ورودی به شبکه باشد. به‌منظور عملیاتی شدن این هدف، شبکه باید بتواند از میان چندین عنوان اتهامی آموزش دیده، یک اتهام را شناسایی و به‌عنوان خروجی ارائه دهد. به نظر می‌رسد در طراحی لایهٔ خروجی، بایستی بر ابر معیار کلاسیفیکیشن چندکلاسه‌ای تمرکز شود؛ زیرا این الگو، دارای طیف گسترده‌ای از کاربردها در تشخیص رقم، طبقه‌بندی برچسب‌گذاری‌ها

۱. توابع غیرخطی به آن دسته از توابع ریاضی گفته می‌شود که رابطهٔ خطی مستقیم بین ورودی و خروجی ایجاد نمی‌کنند.

و تشخیص است (Murphey, 2007, p. 5). برای این منظور، در طراحی ساختار لایهٔ خروجی، در گام نخست، تعداد نورون‌ها باید برابر با تعداد عناوین مجرمانهٔ کلاسه‌بندی شده باشند. سپس، برای این‌که خروجی‌های شبکه به مقادیر احتمالی نرمال‌سازی شده تبدیل شوند تا امکان تفسیر ساده‌تر پیش‌بینی‌ها برای تشخیص عنوان مجرمانه را فراهم سازد، باید از تابع فعال‌سازی (Softmax) بهره گرفته شود؛ زیرا این تابع، مقادیر خروجی شبکه را در لایهٔ خروجی به بازه‌های [۰،۱] تبدیل می‌کند و این امر این امکان را به لایهٔ خروجی می‌دهد تا داده‌های مربوط به هر پروندهٔ قضایی را به‌عنوان احتمال تعلق داده‌های آن پرونده به یکی از عناوین مجرمانه تفسیر کند. برای نمونه، اگر داده‌های پردازش شده در ورودی شبکه با جرائم اموال و مالکیت مرتبط باشند، خروجی لایهٔ Softmax ممکن است احتمال وقوع کلاهبرداری را به میزان ۷۵ درصد، ارتشا را به میزان ۱۴ درصد و اختلاس را به میزان ۱۱ درصد محاسبه کند. در این حالت، شبکه با توجه به بالاترین احتمال (۷۵٪)، عنوان کلاهبرداری را به منزلهٔ خروجی نهایی شبکه انتخاب می‌کند که احتمال خطای تشخیص عنوان صحیح وجود خواهد داشت.

بنابراین به نظر می‌رسد اگر داده‌های یک متن حقوقی به ورودی شبکه به تعداد مختلفی از عناوین مجرمانه تعلق داشته باشند، در این حالت شبکه باید تابع احتمالات را نسبت به هریک از عناوین مجرمانه به‌صورت مستقل با استفاده از تابع Sigmoid محاسبه کند. این تابع احتمال هر عنوان مجرمانه را به‌صورت جداگانه برآورد می‌کند و درنهایت هر عنوانی که احتمال آن از یک مقدار آستانه‌ای بیشتر باشد، به منزلهٔ خروجی در نظر می‌گیرد. برای مثال، اگر محتویات مربوط به یک پروندهٔ کیفری، شامل مفاهیمی مانند جعل اسناد مالی، اختلاس، خیانت در امانت و رشوه باشد، در لایهٔ خروجی، تابع Softmax احتمال وقوع جرم‌های مختلف مانند «جعل»، «خیانت در امانت»، «اختلاس» و... را به‌صورت مستقل محاسبه می‌کند و بالاترین احتمال (مثلاً ۰/۸۵) را، که به‌عنوان «جرم خیانت در امانت» نسبت داده می‌شود، به منزلهٔ خروجی ارائه می‌دهد.

٤. آموزش شبکه برای تشخیص عناوین مجرمانه

پس از معماری مناسب شبکهٔ عصبی، آموزش شبکه برای تشخیص عنوان مجرمانه از اهمیت دوچندان برخوردار است. در واقع، این مرحله زمینهٔ استخراج رابطهٔ میان زمینه و جنبه‌های هدف را از میان مجموعهٔ گسترده‌ای از داده‌های بدون برچسب فراهم می‌سازد (Tong et al., 2022, p. 1). ازاین‌رو، آموزش شبکه باید به‌نحوی صورت پذیرد که تمام لایه‌های پنهان پس از آن‌که به‌صورت تصادفی مقداردهی اولیه شدند، بتوانند متغیرهای ورودی را به آخرین لایهٔ پنهان منتقل کنند، تا وزن لایهٔ خروجی ازطریق راه‌حل‌های حداقل مربعات تخمین زده شود (Yang et al., 2021, p. 151). باوجوداین، موفقیت شبکه در تشخیص عنوان مجرمانه رابطهٔ مستقیم با تعداد مثال‌هایی خواهد داشت که به شبکه ارائه می‌شود؛ زیرا آموزش شبکهٔ عصبی شامل ارائهٔ مثال‌های متعدد از داده‌های برچسب‌گذاری‌شده و سپس به‌روزرسانی وزن‌هاست. این فرایند ازطریق مکانیزم پردازش روبه‌جلو[1] انجام می‌شود، جایی که داده‌ها از ورودی‌های شبکه عبور کرده و ازطریق لایه‌های مختلف به خروجی‌ها منتقل می‌شوند. در هر لایه، داده‌های ورودی (X) با وزن‌های مربوطه (W) ترکیب شده و سپس یک مقدار بایاس (b) به مجموعه اضافه می‌شود که به‌صورت مدل ریاضی زیر خواهد بود:

مدل ریاضی: $\{B+W . X = Z\}$

این مقدار (Z) سپس از یک تابع فعال‌سازی[2] عبور می‌کند تا خروجی نهایی آن لایه تولید شود. این خروجی به لایهٔ بعدی منتقل شده و فرایند در تمام لایه‌ها تکرار می‌شود تا درنهایت شبکه ویژگی‌های کلیدی جرم مدنظر را یاد بگیرد. در پایان، شبکه براساس محاسبات نهایی، عنوان مجرمانهٔ مربوط به داده‌های ورودی را شناسایی و برچسب مناسبی به آن اختصاص می‌دهد. برای مثال، اگر ورودی جمله‌ای مانند «معاون ادارهٔ مالی یک سازمان دولتی بدون رعایت آیین‌نامهٔ معاملاتی دولت اقدام به خرید کالا کرده است» باشد، شبکه پس از پردازش‌های متوالی ممکن است به عنوان مجرمانهٔ مرتبط با «تبانی در معاملات دولتی» (موضوع ماده واحدهٔ قانون مجازات تبانی در معاملات دولتی مصوب ۱۹/۰۳/۱۳۴۸) دست یابد. البته تحقق این هدف، نیازمند برچسب‌گذاری دقیق داده‌ها براساس معیارهای مناسب مانند توصیف عناصر سه‌گانهٔ

1. forward propagation
2. activation function

جرم برای تشخیص الگوهای رفتار هر جرم خواهد بود؛ چراکه روش‌های طبقه‌بندی براساس مدل‌های از پیش آموزش دیده، کمک می‌کند تا شبکه عملکرد خوبی از طبقه‌بندی متن‌های کوتاه (عناوین مجرمانه) از خود نشان دهد (Tong et al., 2022, p. 4).

۵. اعتبارسنجی عناوین اتهامی شناسایی شده

ایدهٔ هوشمندسازی تشخیص عنوان مجرمانه با استفاده از شبکهٔ عصبی NLP رویکردی نوآورانه تلقی می‌شود؛ زیرا این فناوری می‌تواند از داده‌های متنی پرونده‌های قضایی و اسناد مرتبط، با تحلیل زبان طبیعی و الگوهای تکراری در جرائم، به شناسایی عناوین مجرمانه کمک کند. بااین‌حال، ممکن است به‌دلیل پیچیدگی‌های فنی و اصطلاحی متون حقوقی و قضایی و به‌دنبال آن، تفاسیر مختلف از مواد قانون و گستردگی جرائم شبکه در تشخیص عناوین مجرمانه با خطاهای احتمالی روبه‌رو شود. این مسئله زمانی اهمیت بیشتری می‌یابد که مدل‌های شبکهٔ عصبی ممکن باشد دچار سوگیری‌های ذاتی شوند یا تحت‌تأثیر کیفیت داده‌های ورودی قرار بگیرند. برای مثال، در یک پرونده مربوط به جرائم تعزیری، ممکن است شبکهٔ عصبی به‌اشتباه عنوان «سرقت حدی» را برای رفتاری خاص شناسایی کند؛ درحالی‌که رفتار مذکور مشمول قانون «سرقت تعزیری» است. علاوه‌براین، فرض کنید که شبکهٔ عصبی برای شناسایی جرم «جنایات عمدی» موضوع مادهٔ ۲۹۰ قانون مجازات اسلامی مصوب ۱۳۹۲ آموزش داده شده است و یک جمله از پروندهٔ کیفری مانند: «متهم با علم به وضعیت بیماری مجنی علیه اقدام به ایراد ضرب و جرح غیر قابل تحمل کرده است» به‌منزلهٔ دادهٔ آزمایشی به مدل ارائه شود. اگر مدل آموزش داده شده به‌درستی جرم جنایت عمدی را شناسایی کند، عملکرد شبکه مثبت ارزیابی می‌شود؛ اما اگر شبکه به‌اشتباه عنوان مجرمانهٔ دیگری مانند جنایت خطای محض ارائه دهد، این خطا نشان دهند عملکرد غیر قابل دفاع شبکه خواهد بود؛ چراکه این امر می‌تواند به اشتباهات حقوقی و قضایی منجر شود و تأثیرات نامطلوبی بر حقوق متهم، مجنی علیه و حتی سیستم قضایی به دنبال داشته باشد. درخصوص چرایی بروز خطا پیرامون تشخیص عنوان مجرمانه علل و عوامل مختلفی می‌تواند نقش‌آفرین باشند. برای نمونه، ممکن است مدل در تفکیک عنوان جرائم به علت شباهت زبانی دچار خطا شده باشد. برای مثال، به‌دنبال فقدان داده‌های گسترده و عدم تنوع داده‌های قضایی در مدل، شبکه قادر به تفکیک میان جرائم اختلاس و خیانت در امانت نباشد. علاوه‌براین، نبود دقت در برچسب‌گذاری عناوین مجرمانه در مرحلهٔ آموزش داده‌های به مدل، می‌تواند شبکه را در تشخیص عناوین مجرمانه با خطا روبه‌رو سازد؛ ازاین‌رو، اعتبارسنجی خروجی‌های این شبکه ضروری است تا اطمینان حاصل شود که هر عنوان مجرمانه شناسایی شده توسط مدل‌ها هم‌خوان با قوانین و داده‌های پرونده متهم است؛ اما پرسش اساسی این است که براساس چه معیاری‌هایی می‌توان عناوین مجرمانه شناسایی شده توسط شبکهٔ هوش مصنوعی را به‌صورت دقیق و علمی اعتبارسنجی کرد؟ به‌منظور بالا بردن دقت و پایداری شبکه در تشخیص درست عناوین مجرمانه بر پایهٔ مستندات پروندهٔ قضایی متهم و منطبق با قوانین و مقررات، لازم است که در گام نخست عوامل بروز این خطا به‌دقت بررسی شود. یکی از معیارها، اطمینان حاصل کردن از برچسب‌گذاری صحیح داده‌های آموزش داده شده به مدل است. علاوه‌براین، در مرحلهٔ آموزش داده‌ها به مدل، می‌توان داده‌های مربوط به هر یک از جرائم را به دو دسته داده‌های آموزشی و داده‌های آزمایشی تقسیم کرد. براین‌اساس، ابتدا داده‌های آموزشی به میزان ۷۰ درصد به شبکه آموزش داده می‌شود و سپس، خروجی عناوین مجرمانه با استفاده از داده‌های آزمایشی به میزان ۳۰ درصد ارزیابی می‌شود.

البته این امکان وجود دارد که مدل به‌دلیل این‌که عناوین مجرمانه را صرفاً براساس داده‌های آموزشی شناسایی می‌کند در تشخیص عنوان مجرمانه در پرونده‌های کیفری درحال رسیدگی در مراجعی قضایی، با خطای تشخیص مواجه شود. به نظر می‌رسد به‌منظور رفع این چالش و اطمینان از این‌که در عملکرد کلی شبکه، مدل صرفاً براساس داده‌های آموزشی اقدام نمی‌کند، لازم است شبکه با استفاده از مجموعه‌داده‌های جدید ارزیابی فنی شود. برای مثال، جمله‌ای نظیر «تولید کرم‌های آرایشی تقلبی» به مدل داده می‌شود و بررسی می‌شود که آیا شبکه قادر است این موضوع را به‌عنوان «ساخت مواد آرایشی متقلبانه» (موضوع جرائم قانون مواد خوردنی و آشامیدنی و آرایشی مصوب ۱۳۴۶/۰۴/۲۲) شناسایی کند یا خیر؟ همچنین، می‌توان مدل را با استفاده از روش اعتبارسنجی متقابل [۱] ارزیابی کرد. در این روش، داده‌های مربوط به هر جرم به

بخش‌های مختلفی تقسیم می‌شوند. سپس داده‌ها چندین بار آموزش داده شده و آزمایش می‌شوند تا از دقت مدل در تشخیص عناوین مجرمانه اطمینان حاصل شود. به‌عبارت‌دیگر، در این روش داده‌های مربوط به جرم به (K) بخش تقسیم مشود و به تعداد (K) بار مدل برای تشخیص عنوان مجرمانه ارزیابی و بررسی می‌شود. برای مثال، اگر بخواهیم جرم جعل اسناد رسمی (موضوع ماده ۵۳۶ قانون مجازات اسلامی) را به شبکه آموزش دهیم؛ در گام نخست، مجموعه داده‌های مربوط به این جرم شامل متن‌های حقوقی و آرای قضایی گردآوری می‌شود و پس از آن این داده‌ها به‌دقت برچسب‌گذاری می‌شود. سپس داده‌های به دست آمده برای اعتبارسنجی متقابل، به چندین قسمت تقسیم می‌شوند. برای مثال اگر داده‌ها به شش بخش تقسیم شوند، در هر دوره از اعتبارسنجی، یک بخش از داده‌های مربوط به این جرم به‌عنوان مجموعهٔ آزمون و پنج بخش باقی‌مانده به‌عنوان مجموعه‌آموزش استفاده می‌شود. این فرایند شش بار تکرار می‌شود تا تمامی بخش‌ها به‌عنوان یک مجموعه آزمون در یک مرحله استفاده شوند.

٦. چالش‌ها و راهکارهای NLP در مسیر هوشمند سازی تفهیم اتهام

استفاده از فناوری پردازش زبان طبیعی (NLP) در فرایند هوشمندسازی تفهیم اتهام، اگرچه یک تحول فناورانه برای ارتقای عدالت قضایی محسوب می‌شود، دستیابی به این هدف مستلزم شناسایی موانع و ارائهٔ راهکارهای عملی در راستای پیاده‌سازی این فناوری مطابق با نظام قضایی کشور است. به طور خاص، چالش‌های قانونی، اجرایی و ساختاری ازجمله مهم‌ترین موانع در مسیر عملیاتی شدن این فناوری در فرایند تفهیم اتهام به‌شمار می‌روند. از جنبهٔ قانونی، یکی از چالش‌های عمده، نبود چهارچوب‌های حقوقی شفاف و جامع برای استفاده از فناوری‌های هوش مصنوعی در نظام قضایی است. قوانین موجود به‌وضوح جایگاهی برای خروجی‌های سیستم‌های هوشمند به‌عنوان ادله یا ابزار کمکی در تصمیم‌گیری‌های قضایی در نظر نگرفته‌اند، که این امر موجب ایجاد ابهاماتی در پذیرش قانونی این فناوری می‌شود. علاوه‌براین، حفظ حریم خصوصی و امنیت اطلاعات به‌دلیل حساسیت بالای داده‌های حقوقی، از دیگر چالش‌های کلیدی است که نیازمند تدابیر حقوقی دقیق برای پیشگیری از سوءاستفاده و نشت اطلاعات است. همچنین، پیچیدگی‌های فنی و تفسیرناپذیری برخی از خروجی‌ها ممکن است پذیرش این فناوری توسط قضات و وکلا را دشوار سازد. به‌این‌ترتیب، تدوین قوانین و دستورالعمل‌های حقوقی جدید برای استفادهٔ مسئولانه از این فناوری ضروری است. از سوی دیگر، در حوزهٔ چالش‌های ساختاری، محدودیت‌ها و کمبودهای موجود در زیرساخت‌های فناوری اطلاعات، نیروی انسانی متخصص و هماهنگی میان نهادهای قضایی و فناورانه از موانع عمده به‌شمار می‌رود. نبود زیرساخت‌های پیشرفته برای ذخیره و پردازش حجم گستردهٔ داده‌های حقوقی یکی از مشکلات اساسی است که بر دقت و کارایی سیستم‌های NLP تأثیر منفی می‌گذارد. علاوه‌براین، پراکندگی و عدم استانداردسازی داده‌های حقوقی، به‌ویژه اسناد و متون قضایی، موجب کاهش کارایی این فناوری‌ها می‌شود. کمبود متخصصان با دانش میان‌رشته‌ای در زمینه‌های هوش مصنوعی و حقوق نیز یکی دیگر از چالش‌هاست که فرایند توسعه و پیاده‌سازی سیستم‌های پردازش زبان طبیعی را با مشکل مواجه می‌کند. همچنین هماهنگی ناکافی میان دستگاه‌های قضایی و نهادهای فناورانه، روند اجرای پروژه‌های مرتبط را کند کرده و پیچیدگی‌های بیشتری در مقیاس وسیع ایجاد می‌کند. در حوزهٔ اجرایی، چالش‌هایی نظیر مقاومت در برابر تغییرات فناورانه از سوی نیروی انسانی وجود دارد که ناشی از عادت به روش‌های سنتی و نبود اعتماد کافی به سیستم‌های مبتنی بر هوش مصنوعی است. قضات و کارکنان قضایی ممکن است به دقت و شفافیت نتایج تولیدی توسط این سیستم‌ها تردید داشته باشند. علاوه‌براین، هزینه‌های بالای پیاده‌سازی این فناوری، که شامل نیاز به زیرساخت‌های قوی، نرم‌افزارهای پیشرفته و داده‌های با کیفیت بالا است، چالش دیگری است که باید به آن توجه شود. ازاین‌رو، به نظر می رسد استفاده از پردازش زبان طبیعی (NLP) برای احراز عنوان مجرمانه از اوراق پرونده، تحول درخور توجهی در نظام دادرسی ایجاد می‌کند. بااین‌حال، طراحی و پیاده‌سازی یک سیستم کارآمد در این حوزه، مستلزم ایجاد زیرساخت‌هایی نظیر پایگاه داده‌های حقوقی جامع، مدل‌های پردازش متن تخصصی، سیستم‌های استنتاج حقوقی و بسترهای پردازشی قدرتمند است. علاوه‌براین، رعایت ملاحظات امنیتی و حقوقی ازجمله حفظ محرمانگی داده‌ها، کنترل دسترسی و جلوگیری از تصمیم‌گیری خودکار بدون مداخلهٔ انسانی، در اطمینان از صحت و

اعتبار نتایج حاصل از این سیستم نقش اساسی دارد. بااین‌حال، پیشرفت فناوری پردازش زبان طبیعی (NLP) و توسعهٔ مدل‌های یادگیری عمیق در زبان فارسی، مانند مدل‌های مبتنی بر شبکهٔ عصبی عمیق نظیر ParsBERT و DeepSLoT، همراه با دسترسی به منابع حقوقی دیجیتال نظیر سامانهٔ قوانین مجلس شورای اسلامی، قوهٔ قضائیه و دیوان عالی کشور، نیاز به کاهش حجم پرونده‌های قضایی، پیشرفت زیرساخت‌های پردازشی، و امکان به‌کارگیری هوش مصنوعی به‌عنوان ابزار کمک‌قضایی، نشان می‌دهد که استفاده از این فناوری در فرایند تفهیم اتهام در نظام قضایی ایران امکان‌پذیر است؛ به‌ویژه آن‌که دیجیتالی شدن بسیاری از قوانین و آرای قضایی، امکان آموزش مدل‌های NLP برای تحلیل متون حقوقی و استنتاج عنوان مجرمانه از اوراق پرونده را فراهم کرده است. علاوه‌براین، مدل‌های بومی پردازش زبان فارسی با استخراج کلیدواژه‌های کیفری، تطبیق آن‌ها با مواد قانونی، و ارائهٔ پیشنهادهای اولیه به قضات، دقت و سرعت فرایند تفهیم اتهام را افزایش دهند. همچنین، پیشرفت مراکز پردازش ابری داخلی امکان اجرای این مدل‌ها را با حجم بالای داده‌های قضایی امکان‌پذیر کرده است.

نتیجه‌گیری

هوشمندسازی فرایند تفهیم اتهام با استفاده از شبکه‌های عصبی و پردازش زبان طبیعی (NLP)، به‌عنوان یکی از نوآورانه‌ترین کاربردهای هوش مصنوعی در حوزهٔ قضایی، برای بهبود دقت، کارایی و سرعت در شناسایی و تفسیر عناوین مجرمانه از داده‌های متنی، مانند پرونده‌های قضایی و آرای دادگاه‌ها، پتانسیل بالایی دارد. با استفاده از آموزش مدل‌های شبکهٔ عصبی بر مبنای داده‌های حقوقی متنوع و مرتبط، این فناوری قادر است جرائم پیچیده را به‌طور دقیق شناسایی کند و موجب تسهیل فرایندهای قضایی شود. بااین‌حال، اجرای موفقیت‌آمیز این رویکرد مستلزم مدیریت چالش‌هایی نظیر پیچیدگی متون حقوقی، ابهام‌های ناشی از تفاوت‌های تفسیری و زبانی، و سوگیری‌های موجود در داده‌های آموزشی است؛ به‌ویژه بر اهمیت استفاده از داده‌های متنوع و گسترده برای کاهش خطاهای احتمالی و اطمینان از تفکیک دقیق عناوین مجرمانه تأکید می‌شود. همچنین به‌روزرسانی مداوم مدل‌ها مطابق با تغییرات قوانین و مقررات حقوقی ضروری است تا نتایج به دست آمده با قوانین جاری مطابقت داشته باشد. علاوه‌براین، به‌کارگیری این فناوری باید با رعایت ملاحظات حقوقی و اخلاقی انجام شود تا حقوق متهمان و شاکیان در تمامی مراحل حفظ شود و از تصمیم‌گیری‌های ناعادلانه یا مغرضانه جلوگیری شود. ارتقای این سیستم‌ها با توجه به مسائل فنی و حقوقی می‌تواند موجب کاهش بار کاری قضات، افزایش سرعت رسیدگی به پرونده‌ها و ارتقای شفافیت در سیستم‌های قضایی شود. به‌طورکلی، با حل چالش‌های موجود و اتخاذ رویکردهای مناسب، این فناوری تحولی شگرف در سیستم‌های قضایی ایجاد می‌کند و گامی مؤثر در راستای ایجاد یک سیستم قضایی مدرن، کارآمد و عادلانه است.

برای بهره‌برداری بهینه از فناوری هوشمندسازی فرایند تفهیم اتهام با استفاده از شبکه‌های عصبی و پردازش زبان طبیعی (NLP)، پیشنهاد می‌شود مدل‌های شبکهٔ عصبی با داده‌های جامع و متنوع حقوقی آموزش داده شوند تا سوگیری کاهش یافته و دقت در تشخیص عناوین مجرمانه افزایش یابد. ایجاد یک پایگاه دادهٔ استاندارد از متون قضایی، قوانین به‌روزرسانی شده و آرای دادگاه‌ها به بهبود عملکرد سیستم کمک می‌کند و اعتبارسنجی مداوم مدل‌ها ازطریق روش‌هایی نظیر «اعتبارسنجی متقابل» خطاها را به حداقل برساند. در این راستا، توسعهٔ الگوریتم‌های پیشرفته و بومی‌سازی آن‌ها براساس ویژگی‌های خاص نظام حقوقی هر کشور ضروری است. همچنین، آموزش قضات، وکلا و کارشناسان حقوقی درخصوص استفاده از این فناوری‌ها و تفسیر خروجی‌های آن، به همراه تدوین چهارچوب‌های قانونی و اخلاقی برای تضمین رعایت حقوق متهمان و شاکیان، از اهمیت ویژه‌ای برخوردار است. همکاری بین‌المللی میان محققان و نهادهای حقوقی می‌تواند به‌اشتراک‌گذاری دانش و توسعهٔ فناوری‌های کارآمدتر کمک کند. به‌علاوه، سیستم‌های هوشمند باید به‌طور مداوم با تغییرات قوانین و مقررات جدید به‌روزرسانی شوند تا تطابق نتایج با شرایط فعلی حفظ شود. اجرای پروژه‌های پایلوت در مقیاس کوچک پیش از پیاده‌سازی گسترده، امکان شناسایی و اصلاح مشکلات را فراهم می‌آورد و موجب کاهش ریسک می‌شود. درنهایت، حمایت دولت‌ها و نهادهای قانون‌گذار ازطریق سیاست‌گذاری‌های شفاف، تأمین بودجه و تدوین قوانین حمایتی می‌تواند زمینه را برای تحول پایدار در سیستم قضایی فراهم سازد و امکان بهره‌مندی از مزایای این

فناوری نوین را به‌صورت گسترده فراهم کند. براساس یافته‌های این پژوهش، پیشنهاد می‌شود که یک آزمایشگاه هوش مصنوعی تخصصی در حوزهٔ حقوق و قضا تأسیس شود تا الزامات لازم برای بهره‌برداری از فناوری پردازش زبان طبیعی (NLP) فراهم شود. این آزمایشگاه نیازمند تأمین زیرساخت‌های سخت‌افزاری پیشرفته همچون سرورهای پردازشی با پردازنده‌های گرافیکی قدرتمند مانند NVIDIA A100 یا Tesla V100 و حافظهٔ ذخیره‌سازی سریع (SSDNVMe) برای پردازش حجم عظیم داده‌های حقوقی است. همچنین، استفاده از پلتفرم‌های نرم‌افزاری مانند TensorFlowو PyTorch برای توسعهٔ مدل‌هایNLP و بهره‌گیری از پایگاه‌های دادهٔ متنی همچون Elasticsearch و PostgreSQL برای استخراج و تحلیل محتواهای حقوقی ضروری است. به‌علاوه، برای یکپارچه‌سازی پایگاه‌های دادهٔ حقوقی و آرای قضایی، نیاز به تدوین پایگاه داده‌های استاندارد است که شامل متون قوانین، آرای دیوان عالی کشور و پرونده‌های قضایی باشد. در این راستا، حفظ امنیت و حریم خصوصی داده‌ها از اهمیت ویژه‌ای برخوردار است و پیشنهاد می‌شود از فناوری‌هایی مانند رمزگذاری داده‌ها، احراز هویت چندمرحله‌ای و بلاک‌چین برای تأمین امنیت و شفافیت استفاده شود. علاوه‌براین، باید تیم‌های تحقیقاتی مشترک از متخصصان فناوری و حقوق تشکیل شود تا دقت تحلیل‌های مدل‌های هوش مصنوعی افزایش یابد. درنهایت، پیشنهاد می‌شود که چهارچوب‌های قانونی مشخصی برای پذیرش فناوری NLP در فرایند دادرسی تدوین شود تا این فناوری به‌طور رسمی در سیستم قضایی ایران به‌کار گرفته شود. درمجموع، راه‌اندازی این آزمایشگاه می‌تواند به بهینه‌سازی فرایند تفهیم اتهام، کاهش خطاهای انسانی، و افزایش دقت و سرعت در تحلیل پرونده‌های کیفری کمک کند.

منابع

بلوچ فرد، نیلوفر(۱٤۰۱). پردازش زبان طبیعی و تحلیل داده‌های حقوقی در پرونده‌ها و گزارش پزشکی قانونی با استفاده از هوش مصنوعی. تحقیقات نوین میان‌رشته‌ای حقوق، ۱(۲)، ۵۹-۶۴. https://dorl.net/dor/20.1001.1.27833631.1400.2.1.6.2

پاکزاد، عاطفه و آنالویی، مرتضی (۱٤۰۱). ارائه یک رویکرد ترکیبی جدید برای یافتن بردارهای پایه معنادار جهت بازنمایی صریح بردارهای کلمه. رایانش نرم و فناوری اطلاعات، ۱۱(۱)، ۱-۱۷. لینک

جعفرپور، الهام و سیاوشی، سامان (۱۳۹۹). چالش‌های تقنینی و اجرایی تجمیع احکام در فرض تعدد جرم. مطالعات حقوق کیفری و جرم‌شناسی، ۵۰(۲). ۳۹۲-۳۶۹. https://doi.org/10.22059/jqclcs.2021.295723.1515

جلالی شیجانی، فاطمه و شیرزاد، مجید (۱٤۰۰). متن کاوی: مفاهیم و روش‌ها. ترویج علم، ۱۲(۲)، ۱۵۷-۱۷۱. https://doi.org/10.22034/popsci.2022.306089.1130

جعفری پاورسی، حمیده، حریری، نجلا، علیپور حافظی، مهدی، باب الحوائجی، فهیمه و خادمی، مریم (۱۳۹۸). نمایه‌سازی ماشینی مدارک حوزهٔ بازیابی اطلاعات با استفاده از متن‌کاوی در نرم‌افزار رپیدماینر. پژوهشنامهٔ پردازش و مدیریت اطلاعات، ۳۵(۲)، ۳۴۹-۳۷٤. https://doi.org/10.35050/JIPM010.2020.054

رفیع‌پور، ابوالفضل و محسنی، مریم (۱۳۹۸). معلمان ریاضی چگونه از تاریخ تابع در تدریس ریاضی استفاده می‌کنند؟. ریاضی و جامعه، ٤(۲)، ۵۷-۶۶. https://doi.org/10.22108/msci.2020.118430.1333

محمودی جانکی، فیروز و جلیل پیران، علی‌اکبر (۱٤۰۰). تفهیم اتهام در حقوق کیفری ایران و رویهٔ دیوان اروپایی حقوق بشر. مطالعات حقوق تطبیقی، ۱۲(۱)، ۳۱۳-۳۲۹. https://doi.org/10.22059/jcl.2021.314125.634110

Ambarwari, A., Adrian, Q. J., & Herdiyeni, Y. (2020). Analysis of the effect of data scaling on the performance of the machine learning algorithm for plant identification. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *4*(1), 117–122. https://doi.org/10.29207/resti.v4i1.1517

Bashar, A. (2018). Survey on evolving deep learning neural network architectures. *Journal of Artificial Intelligence and Capsule Networks*, *1*(2), 73–82. Link

Brooks, S. K., & Greenberg, N. (2021). Psychological impact of being wrongfully accused of criminal offences: A systematic literature review. *Medicine, Science and the Law*, *61*(1), 44–54. https://doi.org/10.1177/0025802420949069

Chen, H., Wu, L., Chen, J., Lu, W., & Ding, J. (2022). A comparative study of automated legal text classification using random forests and deep learning. *Information Processing & Management*, *59*(2), 102798. https://doi.org/10.1016/j.ipm.2021.102798

Dashti, S. M. S., & Dashti, S. F. (2024). Improving the quality of Persian clinical text with a novel spelling correction system. *BMC Medical Informatics and Decision Making*, *24*(1), 220–237. https://doi.org/10.1186/s12911-024-02613-0

Gharehchopogh, F. S., & Khalifelu, Z. A. (2011). Analysis and evaluation of unstructured data: Text mining versus natural language processing. In *Proceedings of the 5th International Conference on Application of Information and Communication Technologies* (pp. 1-4). https://doi.org/10.1109/ICAICT.2011.6111017

Giménez, M., Palanca, J., & Botti, V. (2019). Semantic-based padding in Convolutional Neural Networks for improving the performance in Natural Language Processing. A case of study in Sentiment Analysis. *Neurocomputing*, *378*, 315–323. https://doi.org/10.1016/j.neucom.2019.08.096

Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. *Organizational Research Methods*, *25*(1), 114–146. https://doi.org/10.1177/109442812097168

Klyuchnikov, N., Titov, A., Kuznetsov, S., Shvets, A., & Ryabinin, M. (2022). NAS-Bench-NLP: Neural architecture search benchmark for natural language processing. *IEEE Access*, *10*, 45736–45747. https://doi.org/10.1109/ACCESS.2022.3169897

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2016). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *1*, 1097–1105. https://doi.org/10.1145/3065386

Lim, S. (2021). Judicial decision-making and explainable artificial intelligence: A reckoning from first principles. *Singapore Academy of Law Journal*, *33*, 280–314. Link

Lombo, X., Oyelade, O. N., & Ezugwu, A. E. (2022). Crime detection and analysis from social media messages using machine learning and natural language processing technique. In O. Gervasi, B. Murgante, S. Misra, A. M. A. C. Rocha, & C. Garau (Eds.), *Computational science and its applications – ICCSA 2022 workshops. ICCSA 2022. Lecture notes in computer science* (Vol. 13381, pp. 7–37). Springer, Cham. https://doi.org/10.1007/978-3-031-10548-7_37

Ma, D., Miniaoui, S., Fen, L., Althubiti, S. A., & Alsenani, T. R. (2023). Intelligent chatbot interaction system capable for sentimental analysis using hybrid machine learning algorithms. *Information Processing & Management*, *60*(5), 103440. https://doi.org/10.1016/j.ipm.2023.103440

Michelbacher, L. (2013). *Multi-word tokenization for natural language processing*. OPUS - Publication Server of the University of Stuttgart. http://dx.doi.org/10.18419/opus-3208

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *Computation and Language*, *1*(7), 1–12. https://doi.org/10.48550/arXiv.1301.3781

Mumcuoğlu, E., Öztürk, C. E., Ozaktas, H. M., & Koç, A. (2021). Natural language processing in law: Prediction of outcomes in the higher courts of Turkey. *Information Processing & Management*, *58*(5), 102684. https://doi.org/10.1016/j.ipm.2021.102684

Murphey, Y. L. (2007). Multi-class pattern classification using neural networks. *Pattern Recognition*, *40*(1), 4–18. https://doi.org/10.1016/j.patcog.2006.04.041

Obanda, S. (2022). *Using Natural Language Processing for Case Brief Generation and Verdict Support in the Kenyan Court System* (Doctoral dissertation, University of Nairobi). Link

Osorio, J., & Beltran, A. (2020). Enhancing the detection of criminal organizations in Mexico using ML and NLP. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-7). IEEE. https://doi.org/10.1109/IJCNN48605.2020.9207039

Robaldo, L., Villata, S., & Wyner, A. (2019). Introduction for artificial intelligence and law: special issue "natural language processing for legal texts". *Artificial Intelligence and Law*, *27*, 113–115. Link

Sharma, V. (2022). A study on data scaling methods for machine learning. *International Journal for Global Academic & Scientific Research*, *1*(1), 31–42. https://doi.org/10.55938/ijgasr.v1i1.4

Singh, P. K., Sharma, S., & Paul, S. (2020). Identifying hidden sentiment in text using deep neural network. In *2nd International Conference on Data, Engineering and Applications (IDEA)* (pp. 1-5). IEEE. https://doi.org/10.1109/IDEA49133.2020.9170726

Tan, P. Y. (2023). The need for artificial intelligence in solving unsolved criminal cases and sentencing in Malaysia. *Asian Journal of Law and Policy*, *3*(3), 89–103. https://doi.org/10.33093/ajlp.2023.7

Tong, J., Wang, Z., & Rui, X. (2022). A multimodel-based deep learning framework for short text multiclass classification with the imbalanced and extremely small data set. *Computational Intelligence and Neuroscience*, *2022*, 7183207. https://doi.org/10.1155/2022/7183207

Wang, D., Li, Y., Jiang, J., Ding, Z., Luo, Z., Jiang, G., Liang, J., & Yang, D. (2024). Tokenization matters! degrading large language models through challenging their tokenization. *arXiv preprint arXiv:2405.17067*. https://doi.org/10.48550/arXiv.2405.17067

Wang, H., He, T., Zou, Z., Shen, S., & Li, Y. (2019). Using case facts to predict accusation based on deep learning. In *2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)* (pp. 133–137). IEEE. https://doi.org/10.1109/QRS-C.2019.00038

Yang, J. (2022). *Research on Security Model Design Based on Computational Network and Natural Language Processing*. Wiley. https://doi.org/10.1155/2022/7191312

Yang, Z., Zhang, H., Sudjianto, A., & Zhang, A. (2021). An effective SteinGLM initialization scheme for training multi-layer feedforward sigmoidal neural networks. *Neural Networks*, *139*, 149-157. https://doi.org/10.1016/j.neunet.2021.02.014

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, *13*(3), 55-75. https://doi.org/10.1109/MCI.2018.2840738

Zhang, W., Wei, W., Wang, W., Jin, L., & Cao, Z. (2021). Reducing BERT computation by padding removal and curriculum learning. In *2021 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)* (pp. 90-92). IEEE. https://doi.org/10.1109/ISPASS51385.2021.00025

Zhong, H., Xiao, C., Tu, C., Zhang, T., Liu, Z., & Sun, M. (2020). How does NLP benefit legal system: A summary of legal artificial intelligence. *arXiv preprint* arXiv:2004.12158. https://doi.org/10.48550/arXiv.2004.12158

[In Persian]

Balouchfard, N. (2022). Natural language processing and analysis of legal data in forensic files and reports using artificial intelligence. *Modern Interdisciplinary Research in Law*, *2*(1), 67-73. https://dorl.net/dor/20.1001.1.27833631.1400.2.1.6.2

Pakzad, A., & Analoui, M. (2022). A novel hybrid approach to finding meaningful basis vectors for explicit representation of word vectors. *Journal of Soft Computing and Information Technology*, *11*(1), 1-17. Link

Jafarpour, E., & Siavashi, S. (2021). Legislative challenges and problems with the enforcement of aggregation of sentences. *Criminal Law and Criminology Studies*, *50*(2), 369-392. https://doi.org/10.22059/jqclcs.2021.295723.1515

Jalali Sheyjani, F., & Shirzad, M. (2021). Text mining: Concepts and methods. *Popularization of Science*, *12*(2), 157-171. https://doi.org/10.22034/popsci.2022.306089.1130

Jafari Powersy, H., Hariri, N., Alipour-Hafezi, M., Bab Al-Hawaiji, F., & Khademi, M. (2020). Machine indexing of documents in the field of information retrieval using text mining in the RapidMiner software. *Iranian Journal of Information Processing and Management*, *35*(2), 349-374. https://doi.org/10.35050/JIPM010.2020.054

Rafiepour, A., & Mohseni, M. (2019). How high school mathematics teachers use the history of functions into their teaching. *Mathematics and Society*, *4*(2), 57-66. https://doi.org/10.22108/msci.2020.118430.1333

Mahmoudi Janaki, F., & Jalilpiran, A. (2021). Arraignment in Iranian criminal law and judicial proceeding of European Court of Human Rights (ECHR). *Comparative Law Review*, *12*(1), 313-329. https://doi.org/10.22059/jcl.2021.314125.634110