

ML-BASED MODELS IN SUPPORT OF SELECTED LEGAL AND ADMINISTRATIVE ACTIVITIES

Monika Łaganowska¹, Maria Mikołajewska², Dariusz Mikołajewski³

¹Politechnika Bydgoska, ul. Kaliskiego 7, 85-796 Bydgoszcz

²Wydział Prawa i Administracji, Uniwersytet Gdański, ul. Bażyńskiego 8, 80-309 Gdańsk

³Wydział Informatyki, Uniwersytet Kazimierza Wielkiego, ul. Kopernika 1, 85-074 Bydgoszcz
e-mail: dariusz.mikolajewski@ukw.edu.pl

Abstract: *This article discusses the application of machine learning (ML) models in improving legal and administrative processes. It highlights how ML techniques such as natural language processing and predictive analytics can automate routine tasks such as document classification, legal research, and case outcome prediction. The authors discuss the benefits of ML-based systems, including increased efficiency, reduced human error, and increased access to justice. Ethical issues are addressed, particularly regarding algorithmic bias, transparency, and accountability in decision-making. Case studies are presented to illustrate the real-world implementation of these technologies in courts and public administration. The article concludes by emphasizing the need for interdisciplinary collaboration and regulatory frameworks to ensure responsible and effective integration of ML in legal domains.*

Słowa kluczowe: *Keywords: Computer science, law, artificial intelligence, machine learning, legal applications, administrative applications.*

MODELE OPARTE NA UCZENIU MASZYNOWYM WSPIERAJĄCE WYBRANE DZIAŁANIA PRAWNE I ADMINISTRACYJNE

Streszczenie: *Artykuł omawia zastosowanie modeli uczenia maszynowego (ML) w ulepszaniu procesów prawnych i administracyjnych. Podkreśla, w jaki sposób techniki ML, takie jak przetwarzanie języka naturalnego i analityka predykcyjna, mogą automatyzować rutynowe zadania, takie jak klasyfikacja dokumentów, badania prawne i przewidywanie wyników spraw. Autorzy omawiają korzyści płynące z systemów opartych na ML, w tym zwiększoną wydajność, zmniejszenie liczby błędów ludzkich i zwiększony dostęp do wymiaru sprawiedliwości. Poruszane są kwestie etyczne, w szczególności dotyczące stroniczości algorytmicznej, przejrzystości i odpowiedzialności w podejmowaniu decyzji. Przedstawiono studia przypadków, aby zilustrować rzeczywiste wdrożenie tych technologii w sądach i administracji publicznej. Artykuł kończy się podkreśleniem potrzeby interdyscyplinarnej współpracy i ram regulacyjnych w celu zapewnienia odpowiedzialnej i skutecznej integracji ML w domenach prawnych.*

Keywords: *Informatyka, prawo, sztuczna inteligencja, uczenie maszynowe, zastosowania prawne, zastosowania w administracji.*

1. Introduction

The genesis of ML-based models in legal and administrative activities stems from the growing need to efficiently manage massive volumes of textual and procedural data. Early attempts to computerize legal tasks focused on rule-based expert systems in the 1970s and 1980s, which were limited by their inflexibility and inability to handle ambiguous or unstructured input. As legal documents and administrative records became increasingly digitized, the

need for more adaptive data-driven tools became apparent. The advent of natural language processing (NLP) enabled machines to analyze legal texts, such as court decisions and statutes, with greater sophistication. In the 2000s, advances in supervised and unsupervised learning facilitated the development of models capable of performing tasks such as legal document classification, contract analysis, and fraud detection. Governments and law firms began to explore predictive analytics to assess case outcomes, optimize resource allocation, and streamline bureaucratic procedures.

The availability of large legal corpora and open government data has accelerated experimentation with ML applications. Over time, pilot projects have shown how automation can support administrative decision-making without completely replacing human judgment. The introduction of deep learning and transformer models has further expanded the capabilities of AI tools in the legal domain, enabling more precise semantic understanding and contextual analysis [1,2]. Initiatives such as predictive policing and automated case triage have illustrated the practical application of ML, although they have also raised concerns about fairness and accountability. Regulators and academia have begun to focus on the ethical, legal, and social implications of ML in public administration [3,4]. This evolution reflects a shift from a purely technological innovation to a multidisciplinary effort to responsibly integrate ML into legal and administrative infrastructures.

The aim of this article is to discuss the possibilities of broader introduction of ML-based computational models into practice in law and administration, especially in Polish conditions.

1.1. Observed research gaps

Despite significant progress, there are still several research gaps in the development and implementation of ML-based models for legal and administrative activities. One of the main gaps is the lack of transparency and explainability in complex models, which hinders trust and accountability in high-stakes decision-making. Data auditing and cybersecurity play an important and underestimated role [5,6]. Many models may still be trained on biased or unrepresentative datasets (e.g., imbalanced across classes), leading to concerns about fairness and the risk of perpetuating existing inequalities (e.g., reflected in real-world historical data). There is still a few research on how to effectively audit and validate ML systems in dynamic legal and administrative environments [7,8]. Another challenge is the insufficient alignment between the design of the technical model and domain-specific legal principles, such as due process and proportionality. ML models often struggle with contextual understanding of legal language, especially in the case of ambiguous or evolving regulations. Interoperability between ML tools and existing legal information systems is also under-researched. In addition, most studies have focused on high-resource legal systems, leaving a gap in applicability to low-resource or multilingual jurisdictions. There is a lack of longitudinal studies assessing the real-world impact and long-term

reliability of ML-supported legal decisions including smart city cases [9,10]. Interdisciplinary collaboration remains limited, which hinders the integration of legal theory, ethics, and eXplainable AI (XAI) in the design and evaluation of such systems including application of IoT, edge computing, federation learning, and cloud computing [11,12].

2. Current concepts

2.1. The impact of regionality on global ML solutions

ML application in law and administration is limited by regional and national regulations on personal data protection, such as the GDPR in the European Union. In countries outside the EU, regulations may be less restrictive or more focused on state control, which affects the scope and manner of implementation of algorithms. Language specificity is a challenge, as ML systems often require large sets of training data in the local language, which are not always available. Cultural differences affect the understanding of justice, which makes it difficult to standardize algorithms used, for example, to assess the risk of recidivism or to allocate benefits. In some countries, public administration is characterized by a low level of digitalization, which limits the possibility of implementing advanced ML systems. Regulations on the transparency of administrative decisions and the requirement to justify them may conflict with the “black box” nature of many ML models. In the EU, the obligation to ensure technological non-discrimination requires thorough testing of algorithms for biases [13,14]. In countries with weaker legal institutions, there is a risk of opaque or illegal use of ML by the administration and extensive security threats [14-16]. International cooperation in e-government requires compatibility of standards, which limits the use of locally adapted, non-standard models. Social resistance resulting from lack of trust in the automation of public decisions can also be a barrier to the implementation of ML in government [17,18].

2.2. Scientific and technological effects of using ML-based models in law and administration

The use of machine learning in law and administration stimulates the development of interdisciplinary research combining computer science, law, ethics and social sciences including automated emotion recognition [19]. Implementing ML in the public sector increases the demand for modern technological solutions, which supports the development of local companies and startups in the AI

industry [20,21]. Creating open public data for model training purposes can support the development of innovative applications and services in the private sector. By investing in ML in administration, Poland can become a regional leader in digitization, which will increase its scientific and technological prestige. New technological challenges, such as the need to ensure transparency and ethics of algorithms, stimulate the development of new methods of eXplainable artificial intelligence (XAI). There are two basic approaches to building models: the first one concerns using only ML models with explainable structure (based on trees or rule models easy to analyze, audit - for financial institutions and other heavily regulated areas), and the second one concerns using more complex models and building their audit on the principle of interrogation (like Post-hoc Model Agnostic). This allows to avoid the "black box" effect. Cooperation between public institutions and universities and research centers can accelerate the transfer of knowledge and the development of digital competences in society. In the long term, effective implementation of ML in the public sector can become an impulse for increasing the innovation of the entire economy, increasing its competitiveness on the international stage.

2.3. Economic effects of using ML-based models in law and administration

Implementing machine learning in law and administration can significantly reduce the state's operating costs by automating routine processes. Increasing the efficiency of public institutions can improve the quality of service to citizens and businesses, supporting economic development. The development of ML technologies in the public sector can stimulate investments in the IT sector and create new jobs in areas such as data science and cybersecurity. However, automation can lead to a reduction in employment in the administration, especially in positions related to data processing and simple official decisions. The use of ML in detecting tax and financial abuses can increase state budget revenues and improve tax collection. Countries that quickly and effectively implement ML in the public sector can gain a competitive advantage in attracting investors and improving the economic climate. However, uneven implementation of these technologies between countries can deepen development differences and lead to greater economic disparities at the regional and global level.

2.4. Social effects of using ML-based models in law and administration

The use of ML in law and administration can increase the efficiency of decision-making processes, shortening the time of handling cases and reducing costs. However, the automation of administrative decisions can lead to a sense of dehumanization and a lack of an individual approach to the citizen, hence it should be ensured that the final decision always belongs to a human specialist. This results from the fact that algorithmic errors or biases in training data can result in unfair decisions, especially towards minority groups. The transparency of decisions made by algorithms is limited, which can reduce public trust in public institutions. This is related to the fact that the use of ML can strengthen control over citizens, especially in authoritarian countries, leading to a deterioration of civil rights (e.g.. GIODO report on the supervised society from 2006). At the same time, properly implemented ML systems can increase the availability of public services, especially in regions with limited administrative resources. These changes can also force the transformation of the labor market in the public sector, leading to a reduction in some positions and the need for new digital competences.

2.5. Possible strategies for implementing and using ML-based models in law and administration

Long-term strategies should be flexible, including in terms of cost, as it requires significant expenditure at the implementation stage, which is not always socially acceptable [21-23]. A possible strategy for implementing ML models in law and administration in Poland may start with pilot projects in selected areas, such as case law analysis or automation of simple administrative decisions. At the same time, it is worth investing in technological infrastructure and creating central repositories of public data, available in a secure and standardized way. A key element of the strategy should be training officials and judges in the basics of ML and algorithmic ethics, which requires an educational program lasting several years. Social campaigns and transparent communication are aimed at building citizens' trust in new technologies and providing them with knowledge of their rights in relation to decisions made by algorithms. Within the first 3–5 years, a legal framework should be created to ensure compliance of the use of ML with the constitution, human rights and European

regulations, such as the AI Act. Implemented models must be tested for transparency, non-discrimination and the ability to explain decisions, which requires cooperation with experts from universities and research centers. In the longer term (5-10 years), the strategy should assume the development of a national AI ecosystem, supporting local research centers (universities, research institutes, production and implementation companies, state and local government units developing guidelines and testing solutions), companies and startups offering innovative solutions for the public sector. Poland can use the implementation of ML in the administration as an impulse to create new jobs in sectors related to data analysis, algorithmic audit and digital security. As a result, transparent, effective and safe use of ML in public institutions can increase the efficiency of the state and also attract foreign technological investments. The overall strategy should be based on the principle of sustainable development, combining innovation with social responsibility and respect for the rights of citizens.

These strategies cannot be implemented in most cases without ensuring access to public data (see e.g. dane.gov.pl) and cooperation between the public and private sectors. Specific information (including public legal acts) can be made available as widely and in the simplest way possible, which will allow for the design and implementation of effective and unbiased algorithms.

It is already possible to create a simple custom GPT model for document processing, what is more: it may be difficult to prevent this, because it speeds up administrative work, while often loading our data to the cloud outside the country. It is necessary to develop and constantly update a comprehensive legal framework regulating the use of AI in administration, conduct cyclical analyses of the impact of AI technology on the rights and freedoms of individuals, and develop ethical standards and best practices for creators and users of AI systems in the area of law and administration.

3. Discussion

We summarized advantages and disadvantages of using ML-based models in law and administration:

1) Advantages:

- Increasing the efficiency and speed of decision-making processes,
- Reducing the costs of administration,

- Automation of routine tasks and relief for employees
- Improving the detection of abuses (e.g. tax, financial),
- Facilitating the analysis of large data sets, e.g. case law,
- Support for the digitalization of public services and increasing the availability of services,
- Potential support for the innovation of the economy and the development of the AI sector,
- Possibility of more effective management of public resources,

2) Disadvantages:

- Data quality dependency and privacy risks,
- Need for large investments in infrastructure and digital skills,
- Risk of bad decisions resulting from errors in data or models,
- Lack of transparency (“black box”) of many ML models,
- Possibility of perpetuating and amplifying biases contained in data,
- Difficulty in ensuring compliance with law (e.g. GDPR, AI Act),
- Risk of limiting individual approach to citizens,
- Low level of social trust and concerns about algorithmic control.

3.1. Limitations

Despite the abundance of publications and research on AI issues, there are still definitional problems, resulting in complex legal issues. ML models in legal contexts still have several limitations that significantly negatively impact their applicability and speed of dissemination. The “black box” nature of some of them (both for the so-called naive user and for AI specialists) makes it difficult to interpret or justify decisions in relation to legal concepts, where literalness is crucial. These models are highly dependent on the quality and representativeness of (real-world) data, which can often be biased, incomplete or outdated. Legal texts are available in a given context, but we do not know how accurately they are interpreted by ML models in this given context. Moreover, ML systems do not have the ability to assess the value that a human making the final interpretative decision brings to the efficiency of analysis and the quality of work and document management. The dynamic and evolutionary nature of regulations, strategies and policies poses a challenge for static models that are not regularly updated. Models trained in one jurisdiction or

used in a specific legal system may not generalize well to other meanings in contexts and procedures. Ethical and legal concerns limit the application of ML to decisions that affect fundamental rights, such as criminal or asylum sentences. Limitations should be highlighted, including the high cost of implementing ML systems, data privacy concerns, and resistance to AI adoption, including without the need for strict regulations and safeguards (e.g., requiring a human expert to make the final decision) that mitigate the potential for bias and undermine public oversight [24,25].

3.2. Directions of further research

Future research on ML-based models for legal and administrative activities should prioritize the development of XAI to ensure transparency and accountability in automated decision-making [26,27]. Researchers need to design models that can be more closely linked to reasoning and legal principles, integrating domain knowledge into the learning process [28]. Addressing algorithmic bias is critical, requiring new techniques to detect bias, mitigate it, and assess fairness in diverse populations. Extending research to underrepresented legal systems and multilingual environments will improve the inclusiveness and applicability of ML tools. Another important direction is the creation of standard reference points and datasets tailored to legal and administrative tasks. Collaboration between lawyers, ethicists, and data scientists should be strengthened to create ethically sound and legally compliant AI systems. Adaptive models that can be updated in response to changes in laws, regulations, or administrative practices are also a promising area of development. Research into the long-term effects of adopting ML in legal institutions is essential to understanding its real-world implications. Integrating human-in-the-loop systems can help combine machine performance with human judgment in sensitive contexts. Developing a regulatory and governance framework for the responsible use of ML in the legal and public sectors will be essential to drive safe, fair, and effective implementation.

4. Conclusions

The rapid introduction of machine learning models in Polish law and administration can be confirmed by the actions of public institutions and shorten the service time. Automation of executive actions relieved by administration employees, them focusing on more complex executive tasks of judgment. ML implementation can also check for problems

with abuse and check the quality of data analysis in the operating system. However, this process was effective, it is necessary to ensure transparent algorithms, network compatibility and social and institutional education. Purposeful and practical use of ML can become an impulse for a digital and economic solution of Poland.

References

1. Li, Q., Peng, H., Li, J., Xia, C., Yang, R., Sun, L., Yu, P.S., He, L. A Survey on Text Classification: From Traditional to Deep Learning. *ACM Trans. Intell. Syst. Technol.* 2022, 13, 1–41.
2. Wan, Z. Text Classification: A Perspective of Deep Learning Methods. *arXiv* 2023, arXiv:2309.13761.
3. Wierczyński, G., Wiewiórowski, W.R. *Informatyka prawnicza. Nowoczesne technologie informacyjne w pracy prawników i administracji publicznej*, wyd. 4, Wolters Kluwer, Warszawa 2016.
4. Ciurak, P., Wierczyński, G. *Nowe technologie w pracy prawnika*, Arche, Gdańsk 2021.
5. Rojszczak, M. *Prawne aspekty systemów sztucznej inteligencji - zarys problemu. Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe*. Wydawnictwo C.H Beck 2019, s. 1-22.
6. Zaidani H, Koulali R, Maizate A, Ouzzif M. Augmentation and Classification of Requests in Moroccan Dialect to Improve Quality of Public Service: A Comparative Study of Algorithms. *Future Internet*. 2025, 17(4), 176.
7. Wierczyński, G., Wiewiórowski, W.R. *Prawne aspekty informatyki, (w:) Informatyka ekonomiczna: teoria i zastosowania*, red. S. Wrycza, J. Maślankowski, Wydawnictwo Naukowe PWN, Warszawa 2019.
8. Wei, F., Keeling, R., Huber-Fliflet, N., Zhang, J., Dabrowski, A., Yang, J., Mao, Q., Qin, H. Empirical Study of LLM Fine-Tuning for Text Classification in Legal Document Review. In *Proceedings of the 2023 IEEE International Conference on Big Data (BigData)*, Sorrento, Italy, 15–18 December 2023; pp. 2786–2792.
9. Sokołowski, M., Wierczyński, G. *Informatyzacja procesu udostępniania informacji prawnych w Austrii i w Polsce – różnice i podobieństwa, (w:) Informatyzacja postępowania sądowego i administracji publicznej*, red. J. Gołaczyński, Warszawa 2010, C.H. Beck, s. 269-278.

10. Wiewiórowski W. Dane osobowe w inteligentnym mieście korzystającym z rozwiązań Internetu rzeczy [w:] Internet rzeczy bezpieczeństwo w smart city [red.:] G. Szpor, C.H.Beck, Warszawa 2015, ss. 24.
11. Nair, A.R., Singh, R.P., Gupta, D., Kumar, P. Evaluating the Impact of Text Data Augmentation on Text Classification Tasks Using DistilBERT. *Procedia Comput. Sci.* 2024, 235, 102–111.
12. Wiewiórowski W. Kwanty informacji o osobie. Prawne aspekty przetwarzania danych o osobach i „objektach” pochodzących z rozproszonych zbiorów [w:] P., Z. Rau, M. Wągrowski [red.:] Nowoczesne systemy łączności i transmisji danych na rzecz bezpieczeństwa. Szanse i zagrożenia, LEX Wolters Kluwer, Warsaw 2013, p. 8.
13. Simmons R. Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System. *SSRN*, 2018, 52 (2), 1067-1118.
14. Wiewiórowski W.R. Profilowanie osób na podstawie ogólnodostępnych danych [w:] A. Mednis [red.:] Prywatność a ekonomia. Ochrona danych osobowych w obrocie gospodarczym, Wyd. WPIA Uniwersytetu Warszawskiego, Warszawa 2013, s. 16
15. Wiewiórowski W.R. Prawne aspekty udostępniania usług administracji publicznej w modelu chmury [w:] G. Szpor [red.:] Internet – Cloud computing. Przetwarzanie w chmurze, C.H.Beck, Warszawa 2013, ss. 38.
16. Wiewiórowski W.R. Prawna ochrona danych biometrycznych w systemach teleinformatycznych pracodawcy. Cele przetwarzania a zakres ochrony [w:] A. Nerka, T. Wyka [red.:] Ochrona danych osobowych podmiotów objętych prawem pracy i prawem ubezpieczeń społecznych. Stan obecny i perspektywy zmian, Wolters Kluwer Polska, Warszawa 2012, s. 10.
17. Wiewiórowski W.R. Automatyzacja rozstrzygnięć i innych czynności w sprawach indywidualnych załatwianych przez organ administracji publicznej [współautor:] G. Sibiga [w:] J. Gołaczyński [red.:] Informatyzacja postępowania sądowego i administracji publicznej, C.H.Beck, Warszawa 2010, s. 13.
18. Gryz, J., Rojszczak M. Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem. *Internet Policy Review* 10 (2), 1-24.
19. Mikołajewska E., Mikołajewski, D. Informatyka afektywna w zastosowaniach cywilnych i wojskowych. *Zeszyty Naukowe/Wyższa Szkoła Oficerska Wojsk Lądowych im. gen. T. Kościuszki*, 2013, 2, 171-184.
20. Veziroğlu, M., Veziroğlu, E., Bucak, İ.Ö. Performance Comparison between Naive Bayes and Machine Learning Algorithms for News Classification. In *Bayesian Inference—Recent Trends*; IntechOpen: London, UK, 2024.
21. Rojek, I., Mroziński, A., Kotlarz, P., Macko, M., Mikołajewski, D. AI-Based Computational Model in Sustainable Transformation of Energy Markets. *Energies* 2023, 16, 8059.
22. Tabany, M.; Gueffal, M. Sentiment Analysis and Fake Amazon Reviews Classification Using SVM Supervised Machine Learning Model. *JAIT* 2024, 15, 49–58.
23. Prokopowicz, P., Mikołajewski, D. Fuzzy approach to computational classification of burnout—Preliminary findings. *Applied Sciences* 2022, 12 (8), 3767.
24. Łagutko, T., Zagajewski, A., Binek, Z., Motykiewicz-Janiak, R., Kończak, A. Wpływ e-commerce na Przemysł 4.0. *Systemy Wspomagania w Inżynierii Produkcji* 2023, 12(2), 99–108.
25. Banaś, D. RegTech jako sposób poprawy efektywności instytucji rynku emerytalnego *Zeszyty Naukowe Politechniki Poznańskiej. Organizacja i Zarządzanie* 2022, 86, 5–19.
26. Bex, F., Prakken, H. Can Predictive Justice Improve the Predictability and Consistency of Judicial Decision-Making? *Frontiers in Artificial Intelligence and Applications*, 2021, 346, 207-214.
27. Ashley, K. A Brief History of the changing roles of case prediction in AI and Law: Law in context *Socio-Legal Journal*, 2019, 36 (1), 93-112.
28. Bhilare, P., Parab, N., Soni, N., Thakur, B. Predicting outcome of judicial cases and analysis using machine learning. *International Research Journal in Engineering Technology*, 2019, 6, 326-330.