

**PEMBANGUNAN MODEL KLASIFIKASI BOT SOSIAL *TWITTER/X*
DENGAN *SEMI SUPERVISED LEARNING* DAN *SUPPORT VECTOR*
*MACHINE***

Disusun dan Diajukan untuk Memenuhi Salah Satu Persyaratan untuk
Memperoleh Gelar Sarjana Komputer Pada Program Studi S1 Rekayasa Perangkat
Lunak



Oleh:

Ardian Hafizh Nurhilman

2007714

PROGRAM STUDI REKAYASA PERANGKAT LUNAK (RPL)

KAMPUS UPI DI CIBIRU

UNIVERSITAS PENDIDIKAN INDONESIA

2024

**PEMBANGUNAN MODEL KLASIFIKASI BOT SOSIAL *TWITTER/X*
DENGAN *SEMI-SUPERVISED LEARNING* DAN *SUPPORT VECTOR
MACHINE***

Oleh:

Ardian Hafizh Nurhilman
2007714

Sebuah skripsi yang diajukan untuk memenuhi salah satu persyaratan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Rekayasa Perangkat Lunak

© Ardian Hafizh Nurhilman
Universitas Pendidikan Indonesia
2024

© Hak Cipta Dilindungi Undang-Undang
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian, dengan dicetak ulang, difoto kopi, atau cara lainnya tanpa izin penulis

HALAMAN PENGESAHAN

ARDIAN HAFIZH NURHILMAN

PEMBANGUNAN MODEL KLASIFIKASI BOT SOSIAL *TWITTER/X*
DENGAN *SEMI-SUPERVISED LEARNING* DAN *SUPPORT VECTOR*
MACHINE

Disetujui dan disahkan oleh:

Pembimbing I



Mochamad Iqbal Ardimansyah, ST., M.Kom.

NIP. 920190219910328101

Pembimbing II



Riditya Muhammad, S.T., MT.

NIP. 920190219920507101

Mengetahui,

Ketua Program Studi Rekayasa Perangkat Lunak



Mochamad Iqbal Ardimansyah, ST., M.Kom.

NIP. 920190219910328101

**HALAMAN PERNYATAAN KEASLIAN SKRIPSI DAN PERNYATAAN
BEBAS PLAGIARISME**

Dengan ini saya menyatakan bahwa skripsi dengan judul “Pembangunan Model Klasifikasi Bot Sosial *Twitter/X* Dengan *Semi-Supervised Learning* Dan *Support Vector Machine*” ini beserta seluruh isinya adalah benar-benar hasil dari karya sendiri. Saya tidak melakukan kegiatan plagiarisme, penjiplakan, ataupun pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Dengan pernyataan ini, saya siap menanggung resiko/sanksi apabila di kemudian hari ditemukan adanya pelanggaran etika keilmuan atau adanya klaim terbukti dari pihak lainnya terhadap keaslian karya saya ini.

Bandung, Juli 2024
Yang membuat pernyataan,



Ardian Hafizh Nurhilman

NIM. 2007714

HALAMAN UCAPAN TERIMA KASIH

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan karunianya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Pembangunan Model Klasifikasi Bot Sosial *Twitter/X* Dengan *Semi-Supervised Learning* Dan *Support Vector Machine*” dengan baik namun tidak terlepas dari kekurangan. Skripsi ini ditulis dan disusun untuk memenuhi salah satu dari syarat untuk memperoleh gelar sarjana dan menyelesaikan studi di Universitas Pendidikan Indonesia pada program studi Rekayasa Perangkat Lunak.

Dalam proses penyusunan dan penulisan skripsi ini ditemui berbagai macam kendala namun dengan pertolongan, bantuan, bimbingan, serta kerja sama dari berbagai pihak kendala tersebut dapat diatasi dan dilalui. Oleh karena itu, perlu bagi penulis untuk mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Orangtua penulis yaitu Bapak Rudiant Komara dan Ibunda Ayuning Budiati yang telah mendoakan, membantu, dan memberikan dukungannya kepada penulis.
2. Bapak Mochamad Iqbal Ardimansyah, ST., M.Kom. selaku dosen pembimbing I yang telah bersedia untuk meluangkan waktunya dan tenaga untuk membimbing penulis dalam proses penelitian dan penyusunan skripsi.
3. Bapak Raditya Muhammad, S.T., MT. selaku dosen pembimbing II yang telah bersedia untuk meluangkan waktunya dan tenaga untuk membimbing penulis dalam proses penelitian dan penyusunan skripsi.
4. Seluruh Rekan Rekayasa Perangkat Lunak 2020 yang telah memberi dukungan selama masa perkuliahan
5. Semua pihak yang telah membantu penulis dalam penelitian dan penyusunan skripsi ini yang tidak dapat disebutkan secara satu persatu.

Penulis menyadari skripsi yang ditulis masih jauh dari kata sempurna. Oleh karena itu, saran dan kritik yang membangun sangat dinantikan oleh penulis demi perbaikan. Penulis mengharapkan bahwa skripsi ini bermanfaat dan memberikan pembelajaran yang berguna bagi penulis dan pembaca khususnya untuk digunakan sebagai sumber ilmu pengetahuan.

Bandung, Juli 2024
Yang membuat pernyataan,



Ardian Hafizh Nurhilman

NIM. 2007714

**PEMBANGUNAN MODEL KLASIFIKASI BOT SOSIAL *TWITTER/X*
DENGAN *SEMI-SUPERVISED LEARNING* DAN *SUPPORT VECTOR
MACHINE***

Oleh:

**Ardian Hafizh Nurhilman – ardian25hilman@upi.edu
2007714**

ABSTRAK

Twitter (atau yang sekarang diberi nama “*X*”) merupakan salah satu jaringan sosial yang paling populer di dunia dan dengan jangkauan yang global. Dengan kepopuleran tersebut salah satu industri yang mencoba untuk memanfaatkan jangkauan global tersebut adalah industri *e-commerce*. Dengan sosial media, sebuah perusahaan *e-commerce* dapat memberikan marketing dan nilai interaksi kepada konsumennya. Namun, pada media sosial seperti *Twitter* sebuah bot sosial dapat mengganggu proses tersebut. Dengan otomasi, sebuah bot sosial dapat mengganggu upaya marketing sebuah perusahaan *e-commerce* di media sosial. Sebuah bot sosial dapat menyebabkan distorsi statistik dan membuat postingan menjadi populer dengan menggunakan otomatisasi. Ditambah lagi, dengan perkembangannya teknologi, sebuah bot sosial lebih sulit dari sebelumnya untuk dibedakan dengan pengguna asli. Berdasarkan masalah tersebut, dibutuhkan cara untuk mengklasifikasikan apakah postingan pada *Twitter* berasal dari pengguna asli atau sebuah bot sosial. Pada penelitian ini, akan dibangun model klasifikasi bot sosial *twitter/x* dengan algoritma *Support Vector Machine*, dan memanfaatkan *semi-supervised learning* dengan proses *pseudo-labelling* untuk meningkatkan kinerja model klasifikasi menggunakan dataset tak berlabel yang dikumpulkan dengan *data crawling*. Setelah model klasifikasi dibangun, dilakukan evaluasi kinerjanya menggunakan *confusion matrix* dan berbagai skenario pengujian. Dari beberapa pengujian, model klasifikasi yang dilatih dengan dataset *pseudo* menghasilkan skor akurasi tertinggi 94% dengan penyetelan parameter SVM $C = 100.0$ dan rasio pembagian data 80:20.

Kata Kunci: Sosial Media, SVM, *Semi-Supervised*, Klasifikasi, *Twitter Bot*

***DEVELOPMENT OF A TWITTER/X SOCIAL BOT CLASSIFICATION
MODEL WITH SEMI-SUPERVISED LEARNING AND SUPPORT VECTOR
MACHINE***

Oleh:

**Ardian Hafizh Nurhilman – ardian25hilman@upi.edu
2007714**

ABSTRACT

Twitter (or what is now called "X") is one of the most popular social networks in the world and with a global reach. With this popularity, one industry that is trying to take advantage of this global reach is the e-commerce industry. With social media, an e-commerce company can provide marketing and interaction value to its consumers. However, on social media such as Twitter a social bot can interfere with this process. With automation, a social bot can disrupt an e-commerce company's marketing efforts on social media. A social bot can cause statistical distortion and make posts popular by using automation. Plus, as technology advances, a social bot is harder than ever to distinguish from a real user. Based on this problem, a way is needed to classify whether posts on Twitter come from real users or social bots. In this research, a Twitter/x social bot classification model will be built using the Support Vector Machine algorithm, and utilizing semi-supervised learning with a pseudo-labelling process to improve the performance of the classification model using unlabeled datasets collected by data crawling. After the classification model is built, its performance is evaluated using a confusion matrix and various test scenarios. From several tests, the classification model trained with the pseudo dataset produced the highest accuracy score of 94% with SVM parameter tuning $C = 100.0$ and a data split ratio of 80:20.

Keywords: Social Media, SVM, Semi-Supervised, Classification, Twitter Bot

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI DAN PERNYATAAN BEBAS PLAGIARISME	iv
HALAMAN UCAPAN TERIMA KASIH	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR RUMUS	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	4
1.4. Manfaat Penelitian	4
1.5. Batasan Masalah	4
1.6. Sistematika Pelaporan Skripsi	4
BAB II KAJIAN PUSTAKA	6
2.1. Media Sosial Sebagai Sumber Informasi	6
2.2. Twitter	6
2.2.1. Twitter Bot	7
2.2.2. Twitter Menjadi X	8
2.3. Data Mining	9
2.4. Bot Sosial	10
2.4.1. Bot Sosial Baik/Jinak dan Bot Sosial Buruk	10
2.5. Klasifikasi	11
2.5.1. Algoritma Support Vector Machine (SVM)	12

2.6. Semi-Supervised Machine Learning	13
2.6.1. Semi-Supervised Support Vector Machine (S3VM).....	14
2.6.2. Pseudo-Labeling.....	15
2.7. State of The Art.....	16
BAB III METODOLOGI PENELITIAN.....	21
3.1. Metode dan Desain Penelitian.....	22
3.1.1. Analyze	22
3.1.2. Design	23
3.1.3. Develop	23
3.1.4. Implement	24
3.1.5. Evaluate.....	24
3.2. Populasi dan Sampel	27
3.2.1. Populasi.....	27
3.2.2. Sampel.....	27
3.3. Instrumen Penelitian	28
BAB IV HASIL DAN PEMBAHASAN	29
4.1. Analyze	29
4.2. Design	32
4.3. Develop	34
4.3.1. Pre-processing	34
4.3.2. Feature Engineering	36
4.3.3. Pembagian Data	42
4.3.4. Pembangunan Model Klasifikasi SVM.....	44
4.3.5. Pseudo-labelling.....	47
4.4. Implement	47
4.5. Evaluate.....	49
4.5.1. Pembahasan Hasil	52
BAB V SIMPULAN DAN SARAN	54
5.1. Simpulan	54
5.2. Saran.....	55

DAFTAR PUSTAKA	56
LAMPIRAN.....	61
Lampiran 1. Tweet Harvest.....	61
Lampiran 2. Pengkodean Tweet-Harvest.....	62
Lampiran 3. Dataset Hasil <i>Twitter Crawling</i>	63
Lampiran 4. Dataset Hasil <i>Cresci-2017</i>	66
Lampiran 5. Dataset <i>Tweet Crawl Feature Engineering</i>	68
Lampiran 6. Dataset <i>Cresci-2017 Feature Engineering</i>	70
Lampiran 7. <i>botometer.osome.iu.edu</i>	72

DAFTAR GAMBAR

Gambar 2.1 Contoh <i>Tweet/Post</i> (twitter.com)	7
Gambar 2.2 Twitter Bot memposting link berbahaya (pandasecurity.com)	8
Gambar 2.3 Proses Data Mining	9
Gambar 2.4 Struktur Metode Klasifikasi	12
Gambar 2.5 SVM untuk Klasifikasi.....	13
Gambar 2.6 Proses <i>Pseudo-labelling</i>	15
Gambar 3.1 Alur Penelitian Menggunakan Model ADDIE.....	21
Gambar 3.2 Nilai Prediksi dan Nilai Aktual	26
Gambar 4.1 Desain Alur Proses Pembangunan Model.....	32
Gambar 4.2 Menghapus Data Kosong atau <i>Null</i>	34
Gambar 4.3 Kolom Dataset (Sebelum <i>Preprocessing</i>).....	35
Gambar 4.4 Kolom Dataset (Setelah <i>Preprocessing</i>)	35
Gambar 4.5 Transformasi Tipe Data.....	36
Gambar 4.6 Frekuensi Fitur <i>Message Based</i>	39
Gambar 4.7 Frekuensi Fitur <i>Special Character Based</i>	40
Gambar 4.8 Frekuensi Fitur <i>Part of Speech Based</i>	41
Gambar 4.9 Frekuensi Fitur <i>Sentiment Based</i>	42
Gambar 4.10 <i>Confusion Matrix</i> Hasil Prediksi Model Klasifikasi	46
Gambar 4.11 <i>Confusion Matrix</i> Hasil Prediksi Model Klasifikasi <i>Pseudo</i>	48
Gambar 4.12 Perbedaan <i>Confusion Matrix</i>	49

DAFTAR TABEL

Tabel 2.1 Persamaan dan Perbedaan Bot Sosial Baik/Jinak dan Bot Sosial Buruk	11
Tabel 2.2 <i>State of The Art</i>	16
Tabel 3.1 <i>Confusion Matrix</i>	25
Tabel 4.1 Atribut Data Hasil <i>Data Crawling</i>	29
Tabel 4.2 Pembagian Data	30
Tabel 4.3 Distribusi Jumlah Akun dan <i>Tweet Dataset Cresci-2017</i>	31
Tabel 4.4 Karakteristik Data	32
Tabel 4.5 <i>Feature Engineering</i>	37
Tabel 4.6 Contoh Hasil <i>Feature Engineering Post Tweet</i>	38
Tabel 4.7 Jumlah Data.....	42
Tabel 4.8 Pengujian Pembagian Data	43
Tabel 4.9 Jumlah Data Latih dan Uji	44
Tabel 4.10 Pengujian <i>Hyperparameter C</i>	45
Tabel 4.11 Hasil Prediksi Model Klasifikasi	46
Tabel 4.12 Jumlah <i>Pseudo-labeled Data</i>	47
Tabel 4.13 Hasil Prediksi Model Klasifikasi <i>Pseudo</i>	48
Tabel 4.14 Evaluasi Kinerja Variasi Pembagian Data	50
Tabel 4.15 Evaluasi Kinerja <i>Hyperparameter C SVM</i>	51

DAFTAR RUMUS

Rumus 3.1 <i>Accuracy</i>	26
Rumus 3.2 <i>Precision</i>	26
Rumus 3.3 <i>Recall</i>	27
Rumus 3.4 <i>F1 Score</i>	27

DAFTAR LAMPIRAN

Lampiran 1. Tweet Harvest.....	61
Lampiran 2. Pengkodean Tweet-Harvest.....	62
Lampiran 3. Dataset Hasil Twitter Crawling.....	63
Lampiran 4. Dataset Hasil <i>Cresci-2017</i>	66
Lampiran 5. Dataset <i>Tweet Crawl Feature Engineering</i>	68
Lampiran 6. Dataset <i>Cresci-2017 Feature Engineering</i>	70
Lampiran 7. <i>botometer.osome.iu.edu</i>	72

DAFTAR PUSTAKA

- Arazzi, M., Ferretti, M., Nicolazzo, S., & Nocera, A. (2023). The role of social media on the evolution of companies: a twitter analysis of streaming service providers. *Online Social Networks and Media*, 36, 100251.
- Ramadhanti, N. S., Kusuma, W. A., & Annisa, A. (2020). Optimasi Data Tidak Seimbang pada Interaksi Drug Target dengan Sampling dan Ensemble Support Vector Machine. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(6), 1221-1230.
- Dixon, S. (2023). Number of Social Media Users Worldwide from 2017 to 2027. Statista. <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>
- Singh, M., & Singh, G. (2018). Impact of social media on e-commerce. *International Journal of Engineering & Technology*, 7(2.30), 21-26.
- Permatasari, A., & Kuswadi, E. (2017). The impact of social media on consumers' purchase intention: A study of ecommerce sites in Jakarta, Indonesia. *Review of Integrative Business and Economics Research*, 6, 321.
- Stieglitz, S., Brachten, F., Ross, B., & Jung, A. K. (2017). Do social bots dream of electric sheep? A categorisation of social media bot accounts. arXiv preprint arXiv:1710.04044.
- De Nicola, R., Petrocchi, M., & Pratelli, M. (2021). On the efficacy of old features for the detection of new bots. *Information Processing & Management*, 58(6), 102685.
- Li, X., Azad, B. A., Rahmati, A., & Nikiforakis, N. (2021, May). Good bot, bad bot: Characterizing automated browsing activity. In *2021 IEEE symposium on security and privacy (sp)* (pp. 1589-1605). IEEE.
- Varol, O., Ferrara, E., Davis, C., Menczer, F., & Flammini, A. (2017, May). Online human-bot interactions: Detection, estimation, and characterization. In *Proceedings of the international AAAI conference on web and social media* (Vol. 11, No. 1, pp. 280-289).

- Godulla, A., Bauer, M., Dietlmeier, J., Lück, A., Matzen, M., & Vaaßen, F. (2021). Good bot vs. bad bot: Opportunities and consequences of using automated software in corporate communications.
- Alsmadi, I., & O'Brien, M. J. (2020). How many bots in Russian troll tweets?. *Information Processing & Management*, 57(6), 102303.
- Wiesenberg, M., & Tench, R. (2020). Deep strategic mediatization: Organizational leaders' knowledge and usage of social bots in an era of disinformation. *International journal of information management*, 51, 102042.
- Bastos, M. T., & Mercea, D. (2019). The Brexit botnet and user-generated hyperpartisan news. *Social science computer review*, 37(1), 38-54.
- Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications*, 10(1), 7.
- Schnebly, J., & Sengupta, S. (2019, January). Random forest twitter bot classifier. In 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 0506-0512). IEEE.
- Wang, A. H. (2010, June). Detecting spam bots in online social networking sites: a machine learning approach. In IFIP Annual Conference on Data and Applications Security and Privacy (pp. 335-342). Springer, Berlin, Heidelberg.
- Kudugunta, S., & Ferrara, E. (2018). Deep neural networks for bot detection. *Information Sciences*, 467, 312-322.
- Sohail, S. S., Khan, M. M., Arsalan, M., Khan, A., Siddiqui, J., Hasan, S. H., & Alam, M. A. (2021). Crawling Twitter data through API: A technical/legal perspective. arXiv preprint arXiv:2105.10724.
- Tian, R., Wang, E., Sivasubramaniam, P., Baskota, S. U., Sharma, A., & Cecchini, M. J. (2024). Ten Years of Pathology on Twitter (X): Landscape and Evolution of Pathology on Twitter From 2012 to 2023. *Archives of Pathology & Laboratory Medicine*.
- Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A. V., & Rong, X. (2015). Data mining for the internet of things: literature review and challenges. *International Journal of Distributed Sensor Networks*, 11(8), 431047.

- Huang, H. Z., Wang, H. K., Li, Y. F., Zhang, L., & Liu, Z. (2015). Support vector machine based estimation of remaining useful life: current research status and future trends. *Journal of Mechanical Science and Technology*, 29(1), 151-163.
- Beskow, D. M., & Carley, K. M. (2018, August). Bot conversations are different: leveraging network metrics for bot detection in twitter. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 825-832). IEEE.
- Ramalingaiah, A., Hussaini, S., & Chaudhari, S. (2021, August). Twitter bot detection using supervised machine learning. In *Journal of Physics: Conference Series* (Vol. 1950, No. 1, p. 012006). IOP Publishing.
- Efthimion, P. G., Payne, S., & Proferes, N. (2018). Supervised machine learning bot detection techniques to identify social twitter bots. *SMU Data Science Review*, 1(2), 5.
- Alarfaj, F. K., Ahmad, H., Khan, H. U., Alomair, A. M., Almusallam, N., & Ahmed, M. (2023). Twitter bot detection using diverse content features and applying machine learning algorithms. *Sustainability*, 15(8), 6662.
- Dwitiyanti, N., Kumala, S. A., & Widiyatun, F. (2020). Using the ADDIE model in development of physics unit conversion application based on Android as learning media. *Formatif: Jurnal Ilmiah Pendidikan MIPA*, 10(2).
- Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). Foundations of data imbalance and solutions for a data democracy. In *data democracy* (pp. 83-106). Academic Press.
- Noor, Shagofah, Omid Tajik, and Jawad Golzar. "Simple random sampling." *International Journal of Education & Language Studies* 1.2 (2022): 78-82.
- Purwanto, P. (2018). Teknik Penyusunan Instrumen Uji Validitas Dan Reliabilitas Penelitian Ekonomi Syariah. *Magelang: Staiapress*.
- Trupthi, M., Pabboju, S., & Narasimha, G. (2017, January). Sentiment analysis on twitter using streaming API. In *2017 IEEE 7th International Advance Computing Conference (IACC)* (pp. 915-919). IEEE.

- Qomari, R. (2009). Teknik penelusuran analisis data kuantitatif dalam penelitian kependidikan. *INSANIA: Jurnal Pemikiran Alternatif Kependidikan*, 14(3), 527-539.
- Chavoshi, N. (2018). *Mining Temporal Activity Patterns On Social Media* (Doctoral dissertation, The University of New Mexico).
- Mbona, I., & Eloff, J. H. (2023). Classifying social media bots as malicious or benign using semi-supervised machine learning. *Journal of Cybersecurity*, 9(1), tyac015.
- Van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. *Machine learning*, 109(2), 373-440.
- Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., & Tesconi, M. (2017, April). The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. In *Proceedings of the 26th international conference on world wide web companion* (pp. 963-972).
- Statista Research Department (2024). Most visited online marketplaces worldwide as of April 2023, by monthly traffic [Infographic]. Statista. <https://www.statista.com/statistics/1155246/leading-online-marketplaces-usa-average-monthly-visits/>
- Muraina, I. (2022). Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts. In *7th international Mardin Artuklu scientific research conference* (pp. 496-504).
- Tharwat, A. (2019). Parameter investigation of support vector machine classifier with kernel functions. *Knowledge and Information Systems*, 61, 1269-1302.
- Bennett, K., & Demiriz, A. (1998). Semi-supervised support vector machines. *Advances in Neural Information processing systems*, 11.
- Ding, S., Zhu, Z., & Zhang, X. (2017). An overview on semi-supervised support vector machine. *Neural Computing and Applications*, 28, 969-978.
- Wang, W., Xi, J., Chong, A., & Li, L. (2017). Driving style classification using a semisupervised support vector machine. *IEEE Transactions on Human-Machine Systems*, 47(5), 650-660.
- Hu, Z., Yang, Z., Hu, X., & Nevatia, R. (2021). Simple: Similar pseudo label exploitation for semi-supervised classification. In *Proceedings of the*

IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15099-15108).

- RÁCZ, A., BAJUSZ, D., & HÉBERGER, K. (2021). Effect of dataset size and train/test split ratios in QSAR/QSPR multiclass classification. *Molecules*, 26(4), 1111.
- LUO, H., MENG, X., ZHAO, Y., & CAI, M. (2023). Rise of social bots: The impact of social bots on public opinion dynamics in public health emergencies from an information ecology perspective. *Telematics and Informatics*, 85, 102051.
- KIM, K. S., YOO-LEE, E., & JOANNA SIN, S. C. (2011). Social media as information source: Undergraduates' use and evaluation behavior. *Proceedings of the American Society for Information Science and Technology*, 48(1), 1-3.
- MURUGANANTHAM, G. (2015). Developing of E-content package by using ADDIE model. *International Journal of Applied Research*, 1(3), 52-54.
- HASNAIN, M., PASHA, M. F., GHANI, I., IMRAN, M., ALZHRANI, M. Y., & BUDIARTO, R. (2020). Evaluating trust prediction and confusion matrix measures for web services ranking. *IEEE Access*, 8, 90847-90861.