







# Marija Stanojevic, Ph.D.

 marija-stanojevic.github.io  
 marijastanojevic

 mstanojevic118@gmail.com  
 marija-stanojevic

 Google Scholar  
 mstanojevic118



## Research Interest

- Multi-modal Learning, Deep Learning, Transfer Learning, Natural Language Processing, Complex and Structured Data, Bioinformatics, Computational Healthcare and Biology




## Employment History

- Aug 2022 – Cambridge Cognition
  - Applied Machine Learning Scientist**, Toronto, ON, Canada.
  - Created multi-modal deep learning architecture, enhancing disease category and severity prediction accuracy by 12% (Transformers, CNN, PyTorch, Docker, AWS, Speech Modeling, Multi-modal Learning, Generative AI).
  - General Chair of Machine Learning for Cognitive and Mental Health Workshop @ AAAI 2024 (Research, Project Lead, Team Lead).
  - Collaborated with pharmaceutical companies on various client projects.
- Jan 2017 – April 2023 Temple University
  - Fellow, Research (RA) and Teaching Assistant (TA)**, Philadelphia, PA, USA.
  - Research Assistant** (Sep - May 2017/18, 2020/21): NSF, NIH, CDC, and IQVIA funded projects (Transformers, RNN, DL, NLP, Graphs, IR, Keras, PyTorch).
  - Teaching Assistant** (2018/19, 2020/21, 2021/22). **Courses:** 1) Data Mining; 2) C and Assembler; 3) Data Structures; **2022 Outstanding Graduate TA Award.**
  - Presidential Fellow** (Jan 2017 - Aug 2020): awarded based on success.
  - Main Organizer** of Mid-Atlantic Student Colloquium on Speech, Language and Learning 2022 (Project Lead, Team Lead).
- Jun – Aug 2021 LinkedIn
  - PhD Machine Learning Engineer Intern**, Philadelphia, PA, USA
  - Proposed and implemented adaptation of neural collaborative filtering with multimodal longitudinal learning (Spark, Scala, Keras, Tensorflow, Dali).
- Jun – Aug 2020 Facebook
  - PhD Machine Learning Engineer Intern**, Philadelphia, PA, USA
  - Designed and implemented a novel multitask multilabel multimodal attention architecture for extreme classification (Python, Presto, PyTorch, Caffe2, DL, Transformers, internal tools).
- Jun – Aug 2019 Facebook
  - PhD Machine Learning Engineer Intern**, Menlo Park, CA, USA
  - Recruiting Science: Improved candidate search by implementing NLP and IR techniques to reduce long tail in skills distribution and by proposing, implementing, and evaluating a novel Siamese-like architecture to embed job descriptions (Python, Presto, PyTorch, Caffe2, DL, NLP, IR, internal tools).
- May – Aug 2018 ADS, Conversant
  - PhD Data Science Intern**, Chicago, IL, USA
  - Pioneered a solution to a large-volume spatio-temporal problem utilizing mean-shift, quick-shift, and hdbscan clustering. Created a proxy to test existing product.
  - Defined evaluation metrics to show potential for implementation into a product (Hadoop, Hive, python, pandas, geo, folium, geopandas, and shapely).

## Employment History (continued)

- Sep 2015 – Jan 2017     **Software Engineer**, Belgrade, Serbia
- Arbor Labs     • Led school performance insight software development, achieving a 30% improvement in efficiency through optimized data cleaning and integration.  
• Reduced costs by 50% by implementing in-house data science and data visualization techniques (PHP, ETL, AWS, R, Python, MySQL, D3.js).

## Education

- 2017 – 2023     **Ph.D., Temple University** in Machine Learning and Data Science.  
Thesis title: *Domain Adaptation Applications to Complex High-Dimensional Target Data*
- 2016 – 2017     **M.Eng., University of Belgrade** in Signal Processing.  
Thesis title: *Determination of the Similarity Between the Scientific Papers Using Machine Learning Methods*
- 2010 – 2016     **B.Eng., University of Belgrade** in Software Engineering.

## Peer-Reviewed Research Publications

### Journal Articles

- 1 Norris, L., **Stanojevic, M.**, & Kendall, P. (in review). Using machine learning to predict treatment outcome in a harmonized dataset of youth anxiety treatments.
- 2 **Stanojevic, M.**, Andjelkovic, J., Kasprowicz, A., Huuki, L. A., Chao, J., Hedges, S. B., ... Obradovic, Z. (2023). Discovering research articles containing evolutionary timetrees by machine learning. *Bioinformatics (Oxford, England)*, 39(1), btado35.
- 3 Andjelkovic, J., Ljubic, B., Abdel Hai, A., **Stanojevic, M.**, Pavlovski, M., Diaz, W., & Obradovic, Z. (2022). Sequential machine learning in prediction of common cancers. *Informatics in Medicine Unlocked*.
- 4 Tarca, A. L., Pataki, B. Á., Romero, R., Sirota, M., Guan, Y., Kutum, R., ... Yu, T. et al. (2021). Crowdsourcing assessment of maternal blood multi-omics for predicting gestational age and preterm birth. *Cell Reports Medicine*, 2(6), 100323.
- 5 Ljubic, B., Hai, A. A., **Stanojevic, M.**, Diaz, W., Polimac, D., Pavlovski, M., & Obradovic, Z. (2020). Predicting complications of diabetes mellitus using advanced machine learning algorithms. *Journal of the American Medical Informatics Association*, 27(9), 1343–1351.

### Conference Proceedings



- 1 Akram, A., Ehghaghi, M., **Stanojevic, M.**, & Novikova, J. (in review). Cross-lingual speaker verification in clinical trials: High performance with no language adaptation. In *Proceedings*.
- 2 Nowenstein, I., **Stanojevic, M.**, Ornlolfsson, G., Jonsdottir, M. K., Simpson, B., Nerin, J. S., ... Curcic, J. (in review). Speech and language biomarkers of neurodegenerative conditions: Developing cross-linguistically valid tools for automatic analysis. In *Proceedings. LREC-COLING 2024 - Joint International Conference on Computational Linguistics, Language Resources and Evaluation*.
- 3 **Stanojevic, M.**, & Novikova, J. (in review). Enhancing multilingual cognitive clinical insights: A transformer-based approach for predictive analysis. In *Proceedings. INTERSPEECH 2024*.
- 4 **Stanojevic, M.** (2024). Machine learning for cognitive and mental health. In *Proceedings. Machine Learning for Cognitive and Mental Health Workshop, AAAI 2024*.

- 5 Ehghaghi, M., **Stanojevic, M.**, Akram, A., & Novikova, J. (2023). Factors affecting the performance of automated speaker verification in alzheimer's disease clinical trials. In *Proceedings. ClinicalNLP Workshop, ACL 2023*.
- 6 Alshehri, J., **Stanojevic, M.**, Dragut, E., & Obradovic, Z. (2022). On label quality in class imbalance setting - a case study. In *Proceedings. 21st International Conference on Machine Learning and Applications, Special Session on Machine Learning for Natural Language Processing, 2022, IEEE*.
- 7 Alshehri, J., **Stanojevic, M.**, Khan, P., Rapp, B., Dragut, E., & Obradovic, Z. (2022). Multilayeret: A unified representation of entities and topics using multilayer graphs. In *Proceedings (pp. 671–687). Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2022. Springer*.
- 8 Diep, B., **Stanojevic, M.**, & Novikova, J. (2022). Multi-modal deep learning system for depression and anxiety detection. In *Proceedings. Empowering Communities: A Participatory Approach to AI for Mental Health, NeurIPS 2022*.
- 9 **Stanojevic, M.**, Norris, L., Kendall, P., & Obradovic, Z. (2022). Predicting anxiety treatment outcomes with machine learning. In *Proceedings. Proc. 21st International Conference on Machine Learning and Applications, Special Session on Machine Learning in Health, 2022, IEEE*.
- 10 Alshehri, J., **Stanojevic, M.**, Dragut, E., & Obradovic, Z. (2021). Stay on topic, please: Aligning user comments to the content of a news article. In *Proceedings (pp. 3–17). European Conference on Information Retrieval, 2021. Springer*.
- 11 Han, C., Cao, X. H., **Stanojevic, M.**, Ghalwash, M., & Obradovic, Z. (2019). Temporal graph regression via structure-aware intrinsic representation learning. In *Proceedings (pp. 360–368). SIAM International Conference on Data Mining, 2019. SIAM*.
- 12 **Stanojevic, M.**, Alshehri, J., Dragut, E. C., & Obradovic, Z. (2019). Biased news data influence on classifying social media posts. In *Proceedings. NewsIR Workshop, SIGIR 2019*.
- 13 **Stanojevic, M.**, Alshehri, J., & Obradovic, Z. (2019). Surveying public opinion using label prediction on social media data. In *Proceedings (pp. 188–195). 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2019. IEEE*.
- 14 Ball, S., **Stanojevic, M.**, Knighton, C., Campbell, W., Thaug, A., Fisher, A., ... Zhou, F. et al. (2018). 2474. early feedback from a pilot of a cognitive computing system to analyze immunization data. In *Proceedings (Vol. 5, S741). Open Forum Infectious Diseases, 2018. Oxford University Press*.
- 15 Brinkley, J., Ball, S., Thaug, A., Campbell, W., Obradovic, Z., **Stanojevic, M.**, ... Fisher, A. (2018). Exploring the metadata of vaccine-related twitter posts: Just how much activity is there and where does it come from? In *Proceedings. Annual Research Meeting, 2018, AcademyHealth*.
- 16 Campogiani, G., Czahajda, R., Mazur, N., & Stanojevic, M. (2014). Involving students in curriculum development. In *Proceedings. European Society for Engineering Education, SEFI Annual Conference, 2014*.
- 17 Stanojevic, M., Martinez, I. S., & Mazur, N. (2014). Virtual internships provided in collaboration among companies and universities-the future of practical development of students. In *Proceedings (pp. 6939–6945). 8th annual International Technology, Education and Development Conference, INTED, 2014. IATED*.

## Books and Chapters



- 1 **Stanojevic, M.**, Alshehri, J., & Obradovic, Z. (2021). High performance computing for understanding natural language. In *Handbook of research on methodologies and applications of supercomputing (pp. 133–144). IGI Global*.

## Skills






- Proficient  • Deep Learning, Transformers, NLP, Research, Multimodality, Transfer Learning, Data Science, Data Mining, Algorithms, Data Structures, Information Retrieval.  
• Python, Keras, PyTorch, C/C++, Java, MySQL, HIVE, Presto.  
• Team, and Project Lead.
- Experienced  • Tensorflow, Hadoop, Bioinformatics, Graphs, CUDA, Docker, Scala, Spark.

## Synergetic Activities







### Awards and Achievements

- 2022  Outstanding Graduate Teaching Assistant Award - Temple University
- 2020-2022  Significant contributor at F31 NIH Fellowship
- 2020  Grace Hopper Celebration (GHC) Student Scholar
- 2017-2020  Temple University Presidential Fellowship
- 2019  Broadening Participation in Data Mining travel & participation award
- 2013  Central European Exchange Program for University Studies (CEEPUS)
- 2012  JoinEUSee (Erasmus Mundus Exchange Program) Scholarship  
 German Academic Exchange Service (DAAD) Summer Course Scholarship
- 2008-2012  Fund for Outstanding Scientific and Art Youth, Ministry of Education, Serbia
- 2010  Award for the top 1% students in Serbia, The Royal Family of Serbia
- 2008  Fund for Young Talents, Ministry of Youth, Serbia: outstanding results award

### Certification


- 2023  **Fundamentals of Accelerated Computing with CUDA C/C++** by NVIDIA
- 2022  **Docker Mastery: With Kubernetes + Swarm from a Docker Captain**
- 2021  **AI for Medicine Specialization** by Deeplearning.ai.  
 **TensorFlow: Advanced Techniques Specialization** by Deeplearning.ai.
- 2019  **Probabilistic Graphical Models Specialization** by Stanford @ Coursera.

### Talks







- Jul, 2023  Multimodal Machine Learning for Healthcare, University of Toronto, Toronto, ON, Canada
- Mar, 2020  Surveying Public Opinion Using Label Prediction on Social Media Data, The 8th Mid-Atlantic Student Colloquium on Speech, Language and Learning
- Oct, 2019  Modeling Scientific Texts, Temple University, Philadelphia, PA
- Apr, 2019  Workshop: Introduction to Artificial Intelligence and Machine Learning, Temple University, Philadelphia, PA
- Aug, 2018  A pilot of a cognitive computing system to analyze immunization data, NSF US-Serbia & West Balkan Data Science Workshop, Belgrade, Serbia
- Jun, 2016  ETL with big data implemented in PHP and SQL, PHP Serbia meetup, Belgrade, Serbia

## Synergetic Activities (continued)






---

- May, 2016  Developing data focused software for insight into education with SCRUM methodology, Faculty of Information Technologies, Metropolitan University, Belgrade, Serbia

### Service and Outreach

- Virtual Chair  ICLR 2021, and ICML 2021
- Associate Editor  Social Network Analysis and Mining (SNAM) journal, Mar 2021 - current
- Reviewer  EMNLP 2022 - current; NAACL 2022-current; ACL ARR 2021-current; ACL 2021-current; EACL 2021-current; ECAI 2023; ECML 2022; Nature Scientific Reports, 2019; NAACL SRW 2022-current; ACL SRW 2021-current; NeurIPS ICBINB 2021-current; NeurIPS DGM4H 2023; Informatics in Medicine Unlocked, 2022; GHC - AI track 2021; Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, 2021; Reproducibility Challenge 2020 - 2021; IMMM 2020; SNAM Journal, 2019; Mary Ann Liebert: Big Data, 2018-2019
- Co-reviewer  KDD 2017
- Mentoring  Five undergraduate and four PhD students
- Main Organiser  9th Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL 2022)
- Co-founder  "Research Mixer" - interdisciplinary research gathering (Feb 2019 - Aug 2020)
- Volunteer  NeurIPS 2020, ACL 2020, ICML 2020, and ICLR 2020
- Research Group Lead  Serbian AI Society, 2021
- Board Member  Technical Workshops Chair at STARS Computing Corps Chapter at Temple University (Spring 2019)
- Instructor  TechGirlz, computer science and machine learning (Feb 2018 - May 2019)
- Soft-skills trainer  Delivered more than 200 hours of soft-skills and technical skills workshops to STEM students across Europe (Board of European Students of Technology - BEST) (2012 - 2016)
- European Management  Board of European Students of Technology (BEST) (2012 - 2013)
- Co-founder  International Science Festival "Science is not Boogeyman" with purpose to promote STEM to students grades 1-12, Nis, Serbia (2008 - 2012)

### Societes

- 2020-now  Member of Association of Computational Linguists (ACL)
- 2019-now  Member of Society for Industrial and Applied Mathematics (SIAM)
- 2018-now  Member of Association for Computing Machinery (ACM)
-  Member of Association for Computing Machinery on Women (ACM-W)
- 2010-2016  Board of European Students of Technology (BEST)

## References

---

Upon request or see LinkedIn