

Vasileios Iosifidis

Machine Learning Scientist

📍 Frankfurt am Main, Germany
@ iosifidis.vasileios.biz@gmail.com
🌐 <https://v-iosifidis.com>
🔗 <https://github.com/iosifidisvasileios>
in <https://www.linkedin.com/in/vasileios-iosifidis>



SUMMARY

Machine Learning/AI specialist with 5+ years of research experience, a solid publication record (25+ papers and articles in AI conferences and journals), and 8+ years of programming proficiency. Skilled in data processing, model design, statistical learning, and data visualization. Expertise in various ML/DL techniques, including Boosting, Bagging, DNNs, Seq2Seq, Clustering, and Random Forests. Proven ability to tackle data challenges, drive innovation, and lead projects successfully.

WORK EXPERIENCE

DB SCHENKER,

Frankfurt, Germany, Senior Data Scientist (Global Data and AI Dep.)

2023, March - Today

- Leading ETA DIRECT project, communicating directly with internal business stakeholders, interpreting requirements and objectives to the right technical solutions, designing and implementing AI system end-to-end.
- Leading Pricing Intelligence project as the Technical Product Owner, coordinating across departments, establishing ETL on internal and external data, designing and deploying deep neural recommendation engine.
- Designed and implemented a Deep Neural Regression ensemble model for ETA LINEHAUL project, enhancing performance by 30.3%. Achieved a 99.4% reduction in training costs and 92% reduction in execution time.
- Designed and implemented (for ETA system) preprocessing methods that corrected 9.4% of the database entries.
- Pioneered a paper club initiative, facilitating in-depth discussions on cutting-edge machine learning methods. Led initial sessions and assisted in group formation for constructive knowledge exchange.
- Implemented Python SOPs (PEP 8, logging, unit testing, versioning, packaging) as a guide for junior team members and other departments.



SCHUFA HOLDING AG,

Wiesbaden, Germany, Data Scientist (Data Quality Dep.)

2021, July - 2023, February

- Collaborated closely with domain experts (curators) to develop and maintain a rule-based engine that efficiently processes over 1 million daily requests.
- Analyzed data to improve algorithmic decisions and create a new set of rules for the rule-based engine, resulting in a 77% reduction in existing expert-annotated errors.
- Designed and implemented Natural Language Processing (NLP) methods/models for data quality, which corrected over 500,000 entries and significantly improved overall accuracy.
- Actively participated in an interdepartmental project, applying NLP techniques that led to a 1.7% correction of overall consumer entries and an 81% reduction in execution time (multi-threading).
- Utilized storage (Big-) data to develop and implement internal tools for system evaluation and notifications, reducing error analysis time and effort to just a few minutes per case.
- Generated a comprehensive dataset by combining various data sources to recalibrate system parameters, resulting in over 3,000 consumer corrections and improved accuracy.



L3S RESEARCH CENTER,

Hannover, Germany, Researcher

2016, June - 2021, June

- Conducted extensive research on topics such as supervised learning, semi-supervised learning, online learning, text classification, class-imbalance, fairness-aware learning, and deep learning resulting in 18 conference and 6 journal publications.
- Led as Project Manager for the NoBIAS ITN project, a Marie Skłodowska-Curie Actions initiative, overseeing a team of 15 Early Research Scientists across 12 organizations in Europe for 2 years.
- Monitored the progress and quality of the NoBIAS project, ensured adherence to EU regulations, managed issue resolution, and implemented corrective action as necessary. Maintained communication with the European Commission regarding project scope, schedule, and budget.
- Supervised B.Sc., M.Sc., and Ph.D. candidates in their academic pursuits.
- Served as a Teaching Assistant for courses such as Artificial Intelligence, Advanced Topics in Data Mining, Data Mining 1, and Data Mining 2.



SENSORFLARE,
Patra, Greece, Full-Stack Developer

2014, December - 2016, April



- Managed and updated the cloud infrastructure, providing access to over 1,000 daily users with minimal downtime and efficient maintenance.
- Successfully integrated multiple IoT APIs, including FitBit, Hue, Insteon, and others into our platform, resulting in an 81% increase in daily traffic and expanded functionality.
- Oversaw the migration of processing functionalities, such as Storm/Zookeeper/Yarn, to AWS and Microsoft Azure to scale our processing capacity, reducing response times by 46% and improving overall efficiency.

RELEVANT SKILLS

Programming Languages: Bash Java OOP Python Scala
Operating Systems: Linux Windows
Big Data/Cloud: Databricks Hadoop Azure ML Studio PySpark
Databases: MySQL Postgres Influx Redis
Collaboration Platforms: Confluence GitLab Jira MS Teams Skype for Business Slack
ML-Frameworks: pytorch pandas scikit-learn hugging-face spacy mlflow
Other: Agile (Kanban, Scrum) MS Office Power BI Public Speaking Team Player

EDUCATION



PH.D. IN MACHINE LEARNING,
Leibniz University of Hannover, Hannover, Germany

2016, June- 2020, July

Faculty of Electrical Engineering and Computer Science, Germany.
Supervisor: Prof. Dr. Eirini Ntoutsis @ ntoutsis@kbs.uni-hannover.de
Thesis title: Semi-supervised learning and fairness-aware learning under class imbalance.
Grade: Summa Cum Laude



M.SC. IN SOFTWARE ENGINEERING,
University of Patras, Patra, Greece

2014, Sept. - 2016, June

Computer Engineering and Informatics Department, Greece.
Supervisor: Prof. Dr. Christos Makris @ makri@ceid.upatras.gr
Thesis title: A Fully Persistent Encryption Binary Tree For Range Queries.
Grade: 9.25/10 (Excellent)






DIPLOMA IN COMPUTER ENGINEERING,
University of Patras, Patra, Greece

2009, Sept.- 2014, August

Computer Engineering and Informatics Department, Greece.
Supervisor: Prof. Dr. Christos Makris @ makri@ceid.upatras.gr
Thesis title: Compressing Inverted Files using Modified LZW.
Grade: 7.15/10 (Very Good)

LANGUAGES

 **Greek**, Native Speaker
 **English**, Professional Level
 **German**, Conversational Level (B2 Telc)



SELECTED LIST OF CERTIFICATES

- Microsoft Certified: Azure Data Scientist Associate
- Structuring Machine Learning Projects
- Software Design and Architecture
- Deep Learning Specialization
- Teamwork Skills: Communicating Effectively in Groups
- Agile with Atlassian Jira
- Conflict Management

SUPERVISED STUDENTS

PhD candidate: Discrimination and bias through algorithms and data, 2021

Master Thesis: Discrimination-Aware Learning in Data Streams, 2018

Master Thesis: Sentiment Analysis with Deep Learning, 2018

Research Project: Generator for Discriminatory Streams, 2018

Research Project: Lexicon-based Approaches for Sentiment Analysis in Twitter, 2017

MEMBER OF PC COMMITTEE

International Journal of Data Science and Analytics, 2024

Expert Systems With Applications (Journal), 2024

AAAI Conference on Artificial Intelligence, 2023

Data Mining and Knowledge Discovery (Journal), Special issue on Bias and Fairness in AI, 2021

European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2020

International Conference on Information and Knowledge Management, 2020

Conference on Hypertext and Social Media, 2020

SELECTED LIST OF PUBLICATIONS

Conference Proceedings

- Roy, Arjun, **Iosifidis, Vasileios**, and Eirini Ntoutsis (2022). "Multi-fairness under class-imbalance". In: *Discovery Science: 25th International Conference, DS 2022, Montpellier, France, October 10–12, 2022, Proceedings*. Springer, pp. 286–301.
- Hu, Hongxin, **Iosifidis, Vasileios**, Wentong Liao, Hang Zhang, Michael YingYang, Eirini Ntoutsis, and Bodo Rosenhahn (2020). "FairNN-Conjoint Learning of Fair Representations for Fair Decisions". In: *23rd International Conference on Discovery Science*.
- **Iosifidis, Vasileios** and Eirini Ntoutsis (2020). "FABBOO - Online Fairness-aware Learning under Class Imbalance". In: *23rd International Conference on Discovery Science*.
- **Iosifidis, Vasileios**, Besnik Fetahu, and Eirini Ntoutsis (2019). "FAE: A Fairness-Aware Ensemble Framework". In: *IEEE International Conference on Big Data, (IEEE BigData) 2019, Los Angeles, USA, December 9-12*. Vol. abs/2002.00695.
- **Iosifidis, Vasileios** and Eirini Ntoutsis (2019a). "AdaFair: Cumulative Fairness Adaptive Boosting". In: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*, pp. 781–790.
- **Iosifidis, Vasileios**, Thi Ngoc Han Tran, and Eirini Ntoutsis (2019). "Fairness-Enhancing Interventions in Stream Classification". In: *Database and Expert Systems Applications - 30th International Conference, DEXA 2019, Linz, Austria, August 26-29, 2019, Proceedings, Part I*, pp. 261–276.
- Fafalios, Pavlos, **Iosifidis, Vasileios**, Eirini Ntoutsis, and Stefan Dietze (2018). "Tweetskb: A public and large-scale rdf corpus of annotated tweets". In: *European Semantic Web Conference*. Springer, pp. 177–190.
- **Iosifidis, Vasileios** and Eirini Ntoutsis (2018). "Dealing with bias via data augmentation in supervised learning scenarios". In: *Jo Bates Paul D. Clough Robert Jäschke*, p. 24.
- Mohapatra, Nilamadhava, **Iosifidis, Vasileios**, Asif Ekbal, Stefan Dietze, and Pavlos Fafalios (2018). "Time-Aware and Corpus-Specific Entity Relatedness". In: vol. abs/1810.10004.
- Fafalios, Pavlos, **Iosifidis, Vasileios**, Kostas Stefanidis, and Eirini Ntoutsis (2017). "Multi-aspect Entity-Centric Analysis of Big Social Media Archives". In: *Research and Advanced Technology for Digital Libraries - 21st International Conference on Theory and Practice of Digital Libraries, TPDL 2017, Thessaloniki, Greece, September 18-21, 2017, Proceedings*, pp. 261–273.
- **Iosifidis, Vasileios** and Eirini Ntoutsis (2017). "Large Scale Sentiment Learning with Limited Labels". In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13 - 17, 2017*, pp. 1823–1832.
- **Iosifidis, Vasileios**, Annina Oelschläger, and Eirini Ntoutsis (2017). "Sentiment Classification over Opinionated Data Streams Through Informed Model Adaptation". In: *21st International Conference on Theory and Practice of Digital Libraries, TPDL*, pp. 369–381.
- **Iosifidis, Vasileios** and Christos Makris (2016). "Compressing Inverted Files using Modified LZW". in: *Proceedings of the 12th International Conference on Web Information Systems and Technologies, WEBIST 2016, Volume 1, Rome, Italy, April 23-25, 2016*, pp. 156–163.

Journal Articles

- **Iosifidis, Vasileios**, Symeon Papadopoulos, Bodo Rosenhahn, and Eirini Ntoutsi (2023). “AdaCC: cumulative cost-sensitive boosting for imbalanced classification”. In: *Knowledge and Information Systems* 65.2, pp. 789–826.
- **Iosifidis, Vasileios**, Arjun Roy, and Eirini Ntoutsi (2022). “Parity-based cumulative fairness-aware boosting”. In: *Knowledge and Information Systems* 64.10, pp. 2737–2770.
- **Iosifidis, Vasileios** (2021). “Über Diskriminierung durch Künstliche Intelligenz”. In: *Ausgezeichnete Informatikdissertationen 2020*.
- **Iosifidis, Vasileios**, Wenbin Zhang, and Eirini Ntoutsi (2021). “Online Fairness-Aware Learning with Imbalanced Data Streams”. In: *CoRR* abs/2108.06231. arXiv: 2108.06231.
- **Iosifidis, Vasileios** and Eirini Ntoutsi (2019b). “Sentiment analysis on big sparse data streams with limited labels”. In: *Knowledge and Information Systems*, pp. 1–40.
- Fafalios, Pavlos, **Iosifidis, Vasileios**, Kostas Stefanidis, and Eirini Ntoutsi (2018). “Tracking the history and evolution of entities: entity-centric temporal analysis of large social media archives”. In: *International Journal on Digital Libraries*, pp. 1–13.