JUSTIN SEBY

✉ justinseby891@gmail.com	☎ +46 762520278	🔗 Portfolio	ⓖ GitHub
in LinkedIn	🎓 Google Scholar	🆔 ORCID	

EDUCATION

Karolinska Institutet, Sweden   September 2024 - Present
Doctor of Philosophy (PhD) in AI for Cancer Precision Medicine   *Stockholm, Sweden*

· Research Focus: Application of advanced machine learning and AI methods on large-scale biological and clinical data for precision medicine in cancer research.
· Developing AI-driven models for biomarker discovery, treatment response prediction, and cancer subtype classification using advanced machine learning techniques.
· Leveraging Explainable AI (XAI) to enhance interpretability and trust in predictive models for clinical decision-making.
· Applying Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) to extract insights from biomedical literature and assist in hypothesis generation.
· Integrating deep learning, probabilistic modeling, and statistical AI approaches to develop robust, data-driven solutions in precision medicine.

University of Oulu, Finland   August 2022 - December 2023
Masters of Science in Computer Science and Engineering   *Oulu, Finland*

· Specialization: Artificial Intelligence (4.45/5 GPA)
· Worked as a student ambassador at the University of Oulu and as a research assistant in AI and ML.
· Courses undertaken: Natural Language Processing and Text Mining, Machine Vision, Data Mining, Machine Learning, etc.
· Thesis Title: ML-based surrogate models for accelerating numerical geophysical simulations

APJ Abdul Kalam Technological University   August 2015 - May 2019
Bachelors of Technology (Computer Science and Engineering   *Kerala, India*

· Graduated with a GPA of 8,29 on a scale of 10 (Top 5)
· Relevant Coursework: Design and Analysis of Algorithm Algorithms, NLP, Artificial Intelligence, Machine Learning, etc.

SKILLS

Natural Language Processing	LLM, Transformer based models like BERT, ROBERTA etc. NLTK, Scrapy, Text Summarization etc. Deep learning models.
Machine Learning Techniques	ML Models, Time Series Analysis.
Programming Languages	Python, Scala, C, SQL, Java and C++ .
Computer Vision	OpenCV, TensorFlow, Keras, PyTorch etc
DevOps Tools	Git, AWS, Docker.
Data Analysis and Visualization	PowerBI, Elastic Search, Data analysis.
Messaging Systems	Kafka.
Performance Testing	JMeter, LoadRunner.
Web Development	Angular/Javascript, HTML, CSS
Frameworks and Libraries	Matplotlib, NumPy. Pandas, Scikit-learn, NLTK, SciPy, TensorFlow Keras etc.
Others	Software documentation skills

SELECTED PUBLICATIONS

Semantic-Driven Focused Crawling Using LASER and FAISS: A Novel Approach for Threat Detection and Improved Information Retrieval
IEEE TrustCom 2023

· Developed a semantic-driven focused crawler for targeted retrieval of Cyber Threat Intelligence (CTI) pages from the internet, combining keyword and link-based strategies with LASER embeddings. Utilized a transformer-based large language model (LLM) classifier to assess the crawler's accuracy and efficiency.

Transformer-based LLMs in Cybersecurity: An in-depth Study on Log Anomaly Detection and Conversational Defense Mechanisms
IEEE International Conference on Big Data 2023

· Explored the potential of chatbots in cybersecurity by integrating anomaly detection features. The chatbot utilizes advanced GPT-3 models and rule-based logic to identify unusual patterns in logs, demonstrating effectiveness in real-world scenarios. The anomaly detection system, combining machine learning and domain expertise, sets a new standard for interactive, anomaly aware conversational agents. Achieved over 99% accuracy in identifying and flagging outliers in log file data, outperforming other Large Language Models (LLMs) such as BERT, DistilBERT, and ALBERT in comparative analysis

CYGENT: A cybersecurity conversational agent with log summarization powered by GPT-3
IEEE International Conference on Artificial Intelligence for Internet of Things 2024

· We developed CYGENT, a conversational agent framework powered by GPT-3.5 turbo model for cybersecurity tasks. CYGENT assists by providing cybersecurity information, analyzing and summarizing uploaded log files, and delivering essential instructions. We achieved a BERTscore of over 97'% in summarizing log files using GPT-3. Our analysis demonstrated that GPT-3 outperformed other Large Language Models, showcasing higher performance.

WORK EXPERIENCE

Research Assistant, DataAI (BISG) June 2023 - Present
University of Oulu *Oulu, Finland*

· Currently working with the University of Oulu's Biomimetics and Intelligent Systems Group on the ADAPTINFA project, funded by the Academy of Finland, to develop data-driven intelligence solutions for multi-hazard mitigation in the Arctic urban environment
· skills: Physics-related Machine learning surrogate modelling, Bayesian optimization, and time series analysis for complex environmental data.

Research Assistant, UBICOMP December 2022 - May 2023
University of Oulu *Oulu, Finland*

· Currently working with the University of Oulu's Biomimetics and Intelligent Systems Group on the ADAPTINFA project, funded by the Academy of Finland, to develop data-driven intelligence solutions for multi-hazard mitigation in the Arctic urban environment
· skills: Physics-related Machine learning surrogate modelling, Bayesian optimization, and time series analysis for complex environmental data.

Software Engineer II July 2021 - December 2021
Sysvine Technologies *Chennai, India*

· Developed codes for ML- based smart captcha service that focus on protecting websites.
· Developed database functions, scripts as part of converting codebase from Ab initio to Azure DataBricks.

System Engineer trainee /Test Engineer August 2019 - March 2021
Infosys *Kerala, India*

- Collaborated with systems architects, consultants and software engineers for development and performance testing processes.
- Handled client calls and participated in requirement analysis discussions. completed 4 projects as per the project requirement.
- Held internal training classes for 10+ employees on JMeter
- Completed Infosys training in advanced Python, data structures, Object oriented programming, Oracle DB and Certified with Advanced level python and java SE8 development.

CERTIFICATIONS

MedTech: AI and Medical Robots	*University of Leeds - Through FutureLearn*
Machine Learning	*Stanford University,Coursera*
The Data Scientist's ToolboxThe Data Scientist's Toolbox	*John Hopkins University ,Coursera*
Mathematics for Machine Learning: Linear Algebra	*Imperial College London - Through Coursera*
AWS Cloud Technical Essentials	*AWS - Through Coursera*
Data Science	*IBM - Through Coursera*

AWARDS AND HONORS

Awarded a University of Oulu scholarship of 14,000 Euro(2022-2024)	*University of Oulu*

LANGUAGE SKILLS

Malayalam	*First language*
English	*Second Language/ Fluent*
Finnish	*Beginner*
Hindi	*Elementary proficiency*

REFERENCES

1. Dr. Jaakko Suutala DSc. (Tech.) : Assistant Professor, Biomimetics and Intelligent Systems Group, University of Oulu *Email: Jaakko.Suutala@oulu.fi*

2. Dr. Satu Tamminen, DSc. (Tech.) : University Lecturer, Biomimetics and Intelligent Systems Group, University of Oulu *Email: satu.tamminen@oulu.fi*