Kristiyan Kamenov Cholakov

kristiyancholakov.com +359886616918 kristiyan-cholakov KristiyanCholakov

Education

Nanyang Technological University, Singapore *Aug 2021 – Jan 2025*
Bachelor of Engineering (B. Eng.) in Computer Science

- Honours (Highest Distinction) — CGPA: 4.59/5.00, Accelerated Bachelor Program (completed in 3.5 years)
- Double Specialization: Artificial Intelligence and Data Science | Minor: Mathematics
- Selected for NTU's URECA program, exclusive for excellent students, conducting research with SOTA ML models.
- Awarded A+ for Final Year Project, earning a nomination for the Nanyang President's Graduate Scholarship.

Experience

TÜV SÜD, Singapore *Jul 2023 – Dec 2023*
Data Engineer

- Designed and implemented the architecture of TÜV SÜD's AI Readiness Analysis app, focusing on scalability, database modeling, and a modular full-stack setup using React.js, Node.js, and Next.js with OAuth authentication.
- Hosted and fine-tuned Meta's Llama 2 on a local server, optimizing the accuracy to generate industry- and client-specific AI evaluations, automating the AI readiness analysis process for both clients and TÜV SÜD's auditors.
- Managed tasks, documentation, and CI/CD pipelines in a trunk-based development workflow with Azure DevOps, ensuring seamless integration, fast iterations, and reliable automated deployments.

Projects

End-to-End Digital Handwritten Examination Platform *Jan 2024 – Nov 2024*
URECA-FYP

- Developed a novel end-to-end platform that seamlessly integrates digital handwritten exam management, layout analysis, content recognition, and automated grading, streamlining the exam workflow.
- Designed and implemented automated STEM grading using LLMs, providing precise and efficient assessments.
- Integrated state-of-the-art models, including PosFormer, TrOCR, and DETR, achieving an impressive 82.19% combined F1 score for handwritten content recognition.
- Fine-tuned DETR on a custom dataset with 10,000 samples, achieving 0.921 F1 in handwritten layout analysis, enhancing the platform's ability to differentiate handwritten text and formulas.

Complexity-Based Handwritten Mathematical Expression Recognition *Sep 2022 – Jun 2023*
URECA

- Combined CoMER (64.39%) and SAN (53.52%) through a complexity-based approach, achieving 65.21% ExpRate.
- Trained a custom classifier for handwritten mathematical expressions by complexity (character count, ink density, and ratio) to select the most suitable recognition model.

Speech Emotion Recognition via Audio-Text Fusion *Feb 2023 – May 2023*
Neural Networks and Deep Learning

- Combined CNN-based text emotion recognition with a MLP for audio features, leveraging OpenAI's Whisper.
- Achieved 71.87% test accuracy, outperforming single-modality approaches and surpassing comparable studies.

ELO Merchant Category Recommendation (Kaggle) *Jan 2024 – Apr 2024*
Machine Learning

- Engineered time-based and behavioral features (aggregate stats, outlier flags, second-order crossovers) to better capture customer purchase patterns and loyalty trends.
- Implemented a dual-model LightGBM approach (outlier vs. non-outlier), improving accuracy by handling extreme cases separately, achieving 3.68232 RMSE (top 6.67% on Kaggle).

Skills

Programming Languages: Python, JavaScript, TypeScript, Java, Kotlin, C, C++, C#, Prolog, SQL

Technologies and Frameworks: TensorFlow, PyTorch, Scikit-learn, Keras, OpenCV, NumPy, Pandas, Matplotlib, Seaborn, XGBoost, LightGBM, Docker, Kubernetes, Git, REST APIs, Flask, React, Next.js, Node.js

Certifications: Google Project Management, Google Business Intelligence, Google IT Automation with Python