Dana Zarezankova

dana.zare@gmail.com | linkedin.com/in/danazarezankova | github.com/dzarezankova

Overview: Engineering student graduating in April 2024 with 2 years of work experience, I want to leverage innovative technology to resolve challenges intrinsic to human experiences across various facets of life.

Skills

Programming Languages: 3+ YoE in Python and MATLAB in both academic coursework and industry work experience.
Machine Learning Tools: Extenstive work in Pandas, NumPy, scikit-learn, PySpark, Matplotlib, Seaborn, Optuna, Tensorflow, and PyTorch.
ML Pipeline Development: Demonstrated ability in all machine learning model development phases, from ideation to data collection, data scraping, cleaning, feature engineering, model selection, tuning, and performance metrics evaluation.

Experience

Machine Learning Intern May – Jul. 2023
Hinge Health *San Francisco, USA*

- Implemented batch processing for SpaCy NLP part of speech tagging to reduce compute time by 60%, patent in progress.
- Fine-tuned PyTorch model to enhance precision and recall for identifying high insurance cost patients by 10%.
- Leveraged NLP APIs like ChatGPT, Jurassic 2, and Claude LLM to create personalized physiotherapy educational articles, utilizing few-shot prompting via LangChain.

Medical Artificial Intelligence Research Assistant Jan. – Apr. 2023
University of Waterloo *Waterloo, Canada*

- Designed and developed an end to end medical record matching system employing random forests, XGBoost, and logistic regression.
- Created a Python library to facilitate the export of scikit-learn models to JSON format, making it easier for cross-platform deployments.
- Authored Python libraries for comprehensive processing of medical record data for enhanced matching accuracy.

Data Scientist May – Aug. 2022
Alife Health *San Francisco, USA*

- Achieved a 30% reduction in the Mean Absolute Error of a complex regression model designed to predict egg retrieval outcomes.
- Conducted a thorough competitive analysis using Monte Carlo simulations to estimate the error margins in competitor models based on their published research.
- Implemented an automated data collection pipeline from the SART online birth rate predictor tool to supplement our existing patient success dataset using Selenium for web scraping.

Education

University of Waterloo Waterloo, ON, Canada
BASc in Honours Systems Design Engineering *Sep. 2019 – Apr. 2024*

- Degree specialization in AI including a capstone engineering project for brain tumour MRI image segmentation.
- Relevant Courses: Data Structures and Algorithms, Linear Signals and Systems, Foundations of AI, Pattern Recognition, Deep Learning, Algorithm Design, Computational Neuroscience, Programming for Performance.
- Served as an academic representative for the class cohort for over 3 years.

Projects

Formula 1 Lap Time Prediction Nov. – Dec. 2023.

- Implemented XGBoost, feedforward, and transformer architectures in PyTorch to compare regression for next lap time, reaching 2.7 second RMSE on tabular and sequential data.
- Pursued hyperparameter tuning using Optuna, increasing model performance by by up to 20%.