Duy Nguyen

Email | LinkedIn | GitHub

Summary

- Proficient in machine learning and AI, leveraging advanced models like SVM, ResNet50, and LLMs.
- Experienced in data analysis, feature engineering, and model evaluation across diverse projects.
- Skilled in Python and frameworks (TensorFlow, PyTorch, Flask, Streamlit) for AI development.
- Strong background in economic data analysis for strategic business decision-making.
- Capable of large-scale data processing using Pandas, NumPy, Neo4j, and Deep Lake.
- Effective collaborator, remote worker, and adept at delivering analytical reports and recommendations.

Education

UC Berkeley College of Engineering and Haas School of Business — Remote
Professional Certificate in Machine Learning and Artificial Intelligence — *Jan 2024 – Present*

Simon Fraser University — Burnaby, BC
Bachelor of Arts in Economics and Data Analysis — *Dec 2022*

Experience

AI Engineer — Mar 2024 – Present
SFU Blueprint

- Transforming Canadian Immigration Consultant using AI knowledge Retrieval Engine. (Shortlisted Top 4 SFU CS Diversity Award)
- Resolved persistent hyperlink functionality issues in AI chatbot, improving user experience and information accessibility.
- Engineered integration of hyperlinks into Neo4j graph database, utilizing vector embedding for efficient retrieval and incorporation.
- Collaborated with front-end engineers to implement parsing solution for AI-generated hyperlinks, bridging gap between backend and frontend systems.
- Demonstrated cross-functional problem-solving skills, enhancing chatbot's practical implementation and overall performance.

Volunteer Software Developer — Jan 2024 – May 2024
Voronoi Health Analytics

- Developed a retrieval augmented generation software that translates doctors' natural language requests into JSON, streamlining access to CT and MRI scan analytics through the DAFs application.
- Eliminated the need for medical professionals to memorize complex codes for report retrieval, potentially saving hours of manual reference time per week.
- Utilized Python and frameworks such as Llama Index and the OpenAI API to develop the software, enhancing efficiency and accuracy in processing natural language to JSON conversion.

Projects

Optimizing Hospital Length of Stay Prediction | *UC Berkeley* — Aug 2024

- **EDA and Visualization:** Conducted in-depth EDA to identify key factors influencing hospital length of stay, including visitor numbers, ward types, and admission types.
- **Multi-Model Approach:** Developed and compared multiple ML models (Gradient Boosting, Random Forest, CatBoost, XGBoost) and a LSTM model, achieving 80.42% accuracy in predicting patient length of stay.
- **Feature Analysis:** Identified key factors impacting stay length: visitor count, ward type, admission deposit, bed grade, and admission type, informing hospital management decisions.
- **Business Impact Assessment:** Demonstrated potential annual savings of $30.4 million and developed a strategic implementation plan for enhancing patient care and operational efficiency.

Optimizing Marketing Campaigns for Customer Subscriptions | *UC Berkeley* Jun 2024

- **EDA and Visualization:** Conducted comprehensive exploratory data analysis, revealing key insights such as higher subscription rates for university-educated customers and a significant class imbalance favoring non-subscribers.
- **Model Evaluation:** Compared four models (Logistic Regression, Decision Tree, KNN, SVM) and identified SVM as the top performer with the highest accuracy and AUC, despite higher computational costs.
- **Feature Importance Analysis:** Utilized permutation feature importance for SVM and Logistic Regression, identifying 'num_emp.var.rate', 'num_duration', and 'num_cons.price.idx' as top features.
- **Strategic Business Recommendations:** Suggested maintaining call durations within 6-8 minutes, reducing campaign contacts, and optimizing re-contact timing to enhance subscription rates.

What drives the price of a car? | *UC Berkeley* Apr 2024

- **Utilized Advanced Analytics:** Applied Lasso and Ridge regression and PCA + KMeans in a used car valuation project to identify key pricing factors, enhancing my ability to refine insurance risk assessment models.
- **Relevant Variables for Insurance Pricing:** Discovered the significant impact of odometer readings and vehicle age on car prices, crucial for setting accurate insurance premiums based on risk profiles.
- **Economic Insights for Strategic Adjustments:** Analyzed how rising fuel prices and changing interest rates affect vehicle sales, providing insights for insurance companies to adjust coverage and premiums in anticipation of market shifts.
- **Targeted Insurance Products:** Identified distinct consumer preferences through segmentation analysis, enabling the design of tailored insurance offerings to meet diverse needs, such as premium packages for luxury vehicles or economical options for older cars.

Embryo Image Classification | *Kaggle* Nov 2023

- Fine-tuned the ResNet50 model for embryo image classification using Python and PyTorch, enhancing accuracy and efficiency in algorithm development and machine learning implementation.
- Achieved a 57.1% accuracy rate and ranked 12th Place on leader board, demonstrating the effectiveness of the deep learning model and optimization techniques.

Publications

Nguyen, D. (2024). *Wave Mechanics: The Key to Generalization in Overparameterized Neural Networks. (In Review)*

- Explores wave properties in the parameter space of overparameterized neural networks.
- Derives conditions for wave function collapse leading to stable generalization.
- Provides theoretical and empirical evidence of damping, oscillations, and implicit regularization.
- Introduces a framework for understanding mini-batch stochastic gradient descent's role in generalization.
- Draws analogies to quantum mechanics while primarily employing classical wave theory.

Nguyen, D. (2024). *Exponential Neuron Networks: A PDE-Inspired Approach to Neural Architecture Design. (In Review)*

- Introduces the Exponential Neuron Network (ENN), a novel architecture inspired by PDE solving techniques (method of characteristics), incorporating exponential activations.
- Proves ENN's universal approximation capability, extending classic results to this new architecture.
- Empirically demonstrates ENN's competitive performance in time series prediction and classification tasks.
- Analyzes ENN's robustness and efficiency in capturing exponential patterns, suggesting promising directions for PDE-inspired neural network designs.

Technical Skills

Languages: Python
Frameworks: Pytorch, Tensorflow, Flask, Streamlit, Neo4j, Deep Lake
Developer Tools: Git, Jupyter Notebook/Lab, VS Code, Visual Studio
Libraries: pandas, NumPy, Matplotlib, seaborn, plotly, pydantic, llamaindex, langchain, transformers, sklearn, pytorch