ANH PHAM

duyanh.phamhoang96@gmail.com • 312-805-0858 • Chicago IL 60605 • Portfolio • LinkedIn

PROFESSIONAL SKILLS

- Computer skills: Excel, Access, Word, PowerPoint, Azure Data Studio, Bloomberg Terminal, SAP
- Programming: Python (NumPy, Pandas, Scipy, Scikit-learn, TensorFlow, PyTorch, BeautifulSoup, NLTK), R (Tidyverse, RANN), PostgreSQL, SAS
- Visualization and Statistical Software: Tableau, Python (Matplotlib, Seaborn, Plotly, Wordcloud), R (Ggplot2), SAS
- Big Data: Hadoop, Hive, Spark
- Analytics, Statistical Analysis, Machine Learning, Financial Modeling, Quantitative Analysis, Security Analysis, Problem Solving, Valuation, Research

PROFESSIONAL EXPERIENCE

FIVERR — Chicago, IL
Freelance Data Analyst — *Jun 2020 - Present*

- Provided data processing, data visualization, and data extracting insights from customer requests using Python, R and SAS
- Applied time series analysis, predictive analysis, and supervised learning (model preparation, classification problems) by applying machine learning algorithms (Logistic and Linear Regression, Decision Trees, Random Forest, SVM) for utilizing dataset
- Strengthened predictions by developing hypermeter optimizations for the dataset and established suggestions for clients to get better predictions in the future

IVYLINE CAPITAL GROUP, LLC. — Lancaster, PA
Apprenticeship Equity Analysis — *Sep 2018 – May 2019*

- Developed complex portfolio management strategies and effectively applied them to security analysis within the investment industry
- Created a successful investment portfolio that used the investment strategy with derivative analytical plan to invest $10 millions paper money on a Think or Swim platform on performed equities: XOM, EOG, GM, BA, and TSLA to gain $1.731 millions in 9 month
- Utilized derivative and statistical analysis effectively to evaluate the return investment in equity options and portfolio management

ANALYTICS PROJECT EXPERIENCE

AMAZON FINE FOOD REVIEW — *Jan 2021*

- Utilized Sparse Matrix by using Python packages (NumPy, Pandas, Scipy) to design recommendation systems based on food item popularity and users' ratings
- Applied text processing with NLTK package and Scikit-learn packages (TfidfVectorizer, CountVectorizer, Gensim) to develop the sentiment analysis that predicts positive and negative reviews by utilizing machine learning technique (Logistic Regression and Bernoulli Naïve Bayes)
- Created clustering model with K-mean to extract top words that impact the sentiment analysis and applied t-SNE to plot those words with Plotly package
- Developed deep learning model to learn the dataset better with ANN and RNN – LSTM by applying TensorFlow.Kera and tunning model to gain higher accuracy scores

AIRPLANE CRASHED — *Nov 2020*

- Utilized Python packages (Pandas, NumPy, Searborn, Matplotlib) to develop a story of the airplane crashed trending based on time, regions, operators, and aircraft
- Built clustering models (K-Mean, Hierarchical, DBSCAN) to cluster airplane crashed and texting on the dataset by utilizing machine learning technique (Scikit-learn) and evaluated the best model based on a loop function of Silhouette Coefficient scores
- Applied dimension reduction (PCA, t-SNE, UMAP) to visualize accuracy results

FASHION MNIST — *Nov 2020*

- Created unsupervised learning/ clustering models (K-Mean, Hierarchical, Gaussian Mixture, DBSCAN) by utilizing machine learning package (Scikit-learn) to cluster the common images
- Evaluated the best model based on a loop function of Silhouette Coefficient scores
- Applied dimension reduction (PCA, t-SNE, LDA, UMAP) to visualize accuracy results
- Developed a 3D plot by applying Plotly

LENDING CLUB'S LOAN DATA FROM 2007 TO 2011 — *Oct 2020*

- Built a model in Python to improve a loan prediction repayment and analysis to the next level by utilizing machine learning technique (Scikit-learn) and by incorporating a strategy for hyperparameter tuning (GridSearchCV) to Radom Forest Classifier and Gradient Boosting Classifier
- Developed data story efficiently and effectively by cleaning out the unnecessary features and pointing out the core variables that have impacts on loan prediction model
- Clarified the best models that fit the dataset with high accuracy scores

EDUCATION

DEPAUL UNIVERSITY, KELLSTADT GRADUATE SCHOOL OF BUSINESS — Chicago, IL
Master of Science in Business Analytics: Data Science — *Jun 2021*

- Dean's Scholarship
- GPA 3.78 / 4.00
- Vice President of Kellstadt Business Analytics Organization

THINKFULL, INC. — Chicago, IL
Data Science Flex Coding Bootcamp — *Feb 2021*

LIBERTY UNIVERSITY, SCHOOL OF BUSINESS — Lynchburg, VA
Bachelor of Science in Business Administration: Finance — *May 2018*