Vishva Shah

shah.vishva5@gmail.com | 773-790-7379 | Chicago, USA

https://vishva-shahh.github.io/ | https://github.com/Vishva-Shahh | https://www.linkedin.com/in/vishva-shah-0201/ | https://medium.com/@Vishva_Shah

Work Experience

Data Science Graduate Assistant | UIC – *Client: Claris Health* | Chicago　　　　　　　　　Aug 2020 - Present
- Detected fraud in opioids oversupply with Auto-encoders and PCA in Python, eventually saving over *~$10 million* in claims.
- Architected end-to-end **LSTM** algorithm for anomaly detection in complex claims data sets to mitigate claim risk by *~25%*.
- Crafted fraud detection system using **machine learning** in healthcare data in PostgreSQL and presenting mitigating measures.

Data Science Research Assistant | UIC | Chicago　　　　　　　　　May 2020 - July 2020
- Applying **NLP** for information retrieval in Python to learn the intent of the input sentence and predict sense of sentence.
- Implemented NLTK Flair model to validate the correctness based on contextual embedding and familiarity coefficients.

Actuarial Data Analyst | ICICI Prudential Life Insurance Company Ltd | India　　　　　　　Jan 2018 - July 2018
- Built **time series** models in R to predict likelihood of an opportunity to convert from early sales, giving *~89%* accuracy.
- Optimized dashboards for the segmentation of customers to target marketing strategies thereby increasing sales by *$8* million.
- **Led** a team of 5 in computing projections, writing, and presenting technical reports outlining risk preventive strategies.

Actuarial Data Analytics Consultant | Reliance Nippon Life Insurance Company Ltd | India　　Nov 2016 - Jan 2018
- Devised **cluster analysis** with K-means segmenting customers, conducted monte-carlo simulations, increasing sales by *~20%.*
- Automated the procedure of extracting massive amounts of results using VBA thereby speeding up the process by *~70%*.
- Performed **multivariate analysis** for financial predictions, cost-benefit analyses, and risk projections for hedging strategies.

Projects

Image classification and captioning using LSTM, Python [GitHub repo]　　　　　　　　　Mar 2020 - Apr 2020
- Performed data augmentation (Rotating, Flipping, Mirroring) increasing training data by *~35%*, reduced overfitting by *~25%.*
- Designed the pretrained CNN model in Pytorch as encoder, used LSTM as decoder, used Flask for Deploying on Heroku.

Variational Auto-Encoders, Python [GitHub repo]　　　　　　　　　Feb 2020 - Mar 2020
- Built a VAE model to generate similar images from dataset, used to increase training data, ELBO loss as optimizer.
- Tuning of hyper-parameters (Batch size, Learning Rate, Momentum) by trial and error and increased training images by *~30%*.

Manipal Hospital Harvard business review Case study, R [GitHub repo]　　　　　　　Sept 2019 - Nov 2019
- Computed statistical measures, imputed ~90% missing values using K-means, increased minority labels by *~40%* with SMOTE.
- Coded random forest and xgboost model, tuned parameters with K-fold, validated results with F1 score of *~87%* and AUC/ROC.

Leadership & Initiatives

Business Analytics Organization [Board of Director] - Leading and managing over 20 MS grads, providing resources for enhancing their skill set, organizing research conferences.

Statistical concepts　　　Published in Medium - Numerical measures, Data types, Probability Distributions

Effects of Social Media　　Published in Medium - Summary of Social Dilemma, Data Privacy

Competencies

Language & Tools　　Python, R, PostgreSQL, PySpark, HTML, Tableau, AWS, MapReduce, Git, Docker, Kubernetes.
Data Science Libraries　　Scikit-Learn, NumPy, SciPy, Pandas, MLlib, OpenCV, ggplot2, rpart, randomforests, ROCR, Caret.
Machine Learning　　Neural Networks (CNN, GAN), Unsupervised Learning (Clustering, PCA), Regression, SVM, K-NN, Decision Trees, Generative Algorithms (Naïve Bayes, LDA).

Education

Master of Science in Business Analytics (STEM), University of Illinois at Chicago, **GPA**: 3.84/4.0　　Expected May 2021
Courses: Machine Learning, Data Mining, Advanced Statistics and analytics, Time Series, Big Data, Network Analysis, ML Deployment, Data Visualization, Multivariate Analytics | **Awarded Full tuition Scholarship.**

Master of Commerce, University of Mumbai (India)　　　　　　　　　Dec 2020

Professional Degree in Actuarial Science, Society of Actuaries (USA) and IFOA (UK)　　　　Apr 2018
Courses: Statistics for risk modelling, Mathematical modelling, Probability, Actuarial mathematics, Predictive analytics, Investment and financial markets, Economics.

Bachelor of Commerce, University of Mumbai (India)　　　　　　　　　Apr 2016