Ishant Virendra Wankhede

✉: ishant30oct@gmail.com ☎: +91 8231025342

Education

B.Tech.	2011-15, Mathematics and Computing **Indian Institute of Technology**, Guwahati	CGPA: **7.38**/10
CLASS XII	2011, **D.A.V Public School, Nerul**	Score: **90.0%**
CLASS X	2009, **D.A.V Public School, Nerul**	Score: **96.6%**

Research Interests

Interpretable Machine Learning, Natural Language Processing, Deep Learning

Achievements

- Speaker at 4th International Conference on Data Management, Analytics and Innovation Delhi'20 on 'Recurrent Neural Networks for Multiclass Multilabel text classification'.
- Secured Rank 81/1020 at AmExpert (Data Science Hackathon) 2019.
- Won 1st Place at Abzooba Data Science Hackathon.
- Awarded 'Employee of the Month' for Sep '19.

Experience

Sr. Software Engineer Data Science *Aug 2015-Present*
Research Associate
Abzooba Inc., Pune
Working on Machine Learning problems in healthcare, retail and insurance sector. Responsible for end-to-end implementation involving feature-engineering, modelling, evaluation and productionization of learning systems, along with infrastructure setup and data visualization across projects:

- Coreference based Pronoun resolution using neuralcoref(SpaCy)
- Simple Question Answering using pre-trained Bi-Directional Attention Flow for Machine Comprehension using ELMO embedding
- Classifier to identify third party subrogation potential. (XGBoost, word2vec, coreNLP, Logistic Regression, Decision Tree)
- Predict product sales from online reviews. (OLS regression, regARIMA, word2vec, coreNLP)
- Fraud detection model for healthcare claims. (Multivariate Gaussian Anomaly detection)
- Classifier to identify future complaints from conversations. (Tf-idf, Affinity Propagation)
- Diabetes prediction model. (L1 regularized Logistic Regression, Decision Tree, ensemble)

Student Trainee *May 2014- July 2014*
Samsung Research Institute, Bangalore

- Worked on enhancing USB-OTG functionalities over Media Transfer Protocol.
- Developed two way master-slave data transfer over OTG for a Phone migration application.

Summer Intern *May 2013- June 2013*
Reliance Industries Ltd, Navi Mumbai

- With the focus being on cost optimisation, logistics and SCM of RIL and export procedures carried out in SEZs, analysed logistics cost for opportunities in cost saving.

- Developed a tool for dynamic suggestion of least cost option for a shipping destination.

Skills

- Application Software: Hadoop Hive, Spark, scikit-learn, BigR, Tableau
- Programming Languages: Python, Java, R, SQL, C/C++, Excel VBA

Projects Undertaken

Finite Difference Methods for Option Pricing
Guide: Prof. Dr. S. Natesan, Dept. Of Mathematics, IIT Guwahati

- Numerical methods for solving Black-Scholes equation for European and Asian options. Used finite differences methods for the discretization of the spatial derivatives, and the Crank-Nicolson scheme for the time derivative.
- Studied the errors in numerical methods for diminishing volatility in options. Used non-uniform mesh for spatial discretisation to minimize error, and compare results with uniform mesh.

Identify third party subrogation potential from unstructured insurance adjusters' notes

- To find the person/party at fault for the accident, nature of the accident, party suffering damages.
- We built a Document2Vector, expanding Word2Vector, defining features for our usecase.
- I worked on capturing and enhancing typed dependencies from dependency tree, building generalized constructs of phrases, to capture cause-and-damage from text.
- I built heuristic rule engine and XGBoost classifier, ensembled with Logistic Regression classifier. Worked on transfer learning, dynamic AIC feature selection.
- Built framework for distributing the application across multiple servers

Predicting product sales explained by online reviews and social media content

- Built knowledge-hierarchy from word2vec entities in order to map online reviews to product attributes.
- Quantifying the text review by sentiment analysis, we were able to show that online reviews are correlated to sales. Further bowdlerizing the claim, we were able to model them, using regARIMA, to predict future sales.

Multi-variate Gaussian fraud detection model for Insurance claims

- Built an anomaly detection model from scratch in SAS a data analytics tool, without any pre-existing functions or libraries; to Identify Fraud, Waste, and Abuse in medical claims.
- After going through a research phase, identified and designed hypotheses, which became important features, such as up-coding, and overbilling.

Diabetes prediction model

- Worked on a statistical model which predicts the probability of being diabetic in the next two years. I ensured deployment, performed testing and built error logging modules to track performance.
- Studied and observed the model and feature engineering, to develop in future. Built a visualization tool on top of model results, to generate actionable insights.

Sampling of posterior of Weibull Distribution parameters
Guide: Dr. Arabin Kumar Dey , Dept. Of Mathematics, IIT-Guwahati

- Generated samples from posterior of modified Weibull distribution using Gibbs sampling and Metropolis-Hastings Algorithm using R and MATLAB.