

Correspondence Address:
Flat A103, Euphoria
Suncity Apartments, Ibbalur
Bellandur, Bengaluru- 560102

Vivek Kumar
vivekkumar531@gmail.com
+91 8870486333

Permanent Address:
Upper Bazar, Mahabir Chowk
Pyada Toli, Ranchi
Ranchi-834001

Education **IIT (BHU) Varanasi** **IDD (CGPA: 7.89/ 10)** *July 2009 - June 2014*

Introduction • I am a **Data Scientist** working in Walmart Labs, Bengaluru
 • I am also mentoring people interested in Data Science on Springboard

Skillset **Technology & Techniques**
 • R, Python, C++, MySQL, Excel, Git, AWS
 • Flask, MongoDB, Docker and Kubernetes
 • Probabilistic Models, Statistics, HMM, CRF
 • Machine Learning - CNN, RNN, LSTM, Random Forest, Gradient Boosting
 • Numpy, Pandas, Scikit-learn, Keras, Tensorflow
 • Data Analysis Libraries: e1071, tm, ggplot2, RWeka, data.table, xgBoost

Experience **Walmart Labs, Bengaluru** Dec2017-Present
 Engineer 3, Data Science
 Working with the Catalog data sciences team as a Data Scientist. Building text classification and sequencing models for extracting attributes from diverse set of product data (images, text.)
 Deployment of Containerized Flask model application on kubernetes cluster.

Altisource Labs, Bengaluru *June 2016-Dec 2017*
 Lead, Data Science
 • **Classification of mortgage documents (NLP)**
 Working on classification of mortgage documents in 300+ classes and separate document boundaries. Used Random forest algorithm, XGBoost, SVM and naive Bayes to successfully classify documents with above 85% accuracy and modified Hidden Markov Model for document separation with above 84% accuracy on very large dataset.

Global Analytics, Chennai *June 2015-June 2016*
 Analytics Scientist
 • **Risk Analytics**
 Developed and utilized the probability of default models for both new and existing customers for sub-prime lending portfolio in UK. Web scraping to find useful variables to reduce data cost. Used a lot of different machine learning algorithms to increase the KS score of the models.
 • **Revenue prediction (forecasting)**
 Developed cash flow models using linear regression and shrinkage methods. Monitored the models on periodic basis to check the performance of existing models.

Springboard April 2016-Present
 Mentor, Data Science Intensive Program
 Mentored several people interested in Data Science, Machine Learning

IBM India Pvt. Ltd., Bengaluru Dec 2014-June 2015
 Associate systems engineer

Achievements • Achieved 99.85%ile in CAT13
 • Cleared IITJEE 2009 with AIR 5328 and AIEEE with AIR 3131