

Sector 125, Noida, India

91764-31675

[swagatadutta.official@gmail.com](mailto:swagatadutta.official@gmail.com)

Skype: swagatadutta.official

# SWAGATA DUTTA

## EXPERIENCE

### **Spicemoney, Noida** – *Data Scientist.*

JULY 2019 – Present

#### **About the company:**

SpiceMoney is a Hyper Local Payments Network offering various services like Cash Deposit, Cash Withdrawal, Balance Inquiry, Bill Payments, Aadhaar Enabled Services, Air Time Recharge, POS Services, Railway Ticketing Services etc. through authorized agents based out of rural marketplace, endeavouring to bridge the gap between the served and the un-served, providing banking services to customers at their doorstep.

#### **Projects Undertaken**

- **CHURN PREDICTION:**
  - Developed churn prediction model that helps as early warning signal of possible usage and category churn.
  - Created report on progress of risk and breakdown of reasons leading to churn – identifying cause behind churn.
  - Created report for tracking impact of churn prediction algorithm.
- **REPORTS AND DASHBOARDS:**
  - Created consumer insights dashboard to track the extent of consumption of our product by the end customers.
  - Created django api for CSP scorecard module.
  - Created daily competition analysis report
  - Developed analysis on our presence and presence of competition by geography and distributors and provided actionables on where we are losing ground.
  - Developed and automated reports and dashboard for actionable intelligence for Customer Lifecycle management team.
  - Created reports to track effectiveness of onboarding of new CSPs and their time to maturity
  - Enhanced analytics data warehouse with more pre-processed data about CSPs, Consumers and Products.
- **CSP LOYALTY REWARDS:** Created grading system and featurization for CSP Loyalty Rewards Program.
- **NEARBY CSP:** Developed Algorithm to track nearest Active CSPs to the end customers, who are more probable to hold cash at their counter for Domestic Money Transfer Service and deployed it as a django API.
- **CONSUMER APP RECOMMENDATION SYSTEM:** Cross-selling of products and estimation of cash availability at consumer end by behaviour tracking, to publish recommendations for advising usage of appropriate product.

### **DataToBiz, Chandigarh** – *Machine Learning Engineer*

NOVEMBER 2018 – JUNE 2019

#### **About the company:**

As a Data Science Solutions provider, DataToBiz provides competitive insights to clients, empowering them in taking data-informed decisions using advanced machine and deep learning technologies.

### **Projects Undertaken**

- **FACE ATTENDANCE SYSTEM:** Developed a face attendance system for classroom scenario, which leverages the state-of-the-art face recognition algorithm 'Arcface' developed by Deepinsight. The system currently works on video feeds. It integrates the implementation of detection, blur correction, exposure correction, recognition and tracking. Platforms used extensively are OpenCV and MXNet with Python.
- **RECOMMENDATION ENGINE:** Developed a Recommendation System for a Telco Client with implicit interactions of transactions. Used Non-Negative Matrix Factorisation with Alternating Least Square loss for collaborative filtering model. Product coverage achieved for the system for top 10 ranked recommendations was more than 60% and Mean Average Precision achieved for the same was 35%.
- **CREDIT RISK MODEL:** Prepared a credit risk model for a Telco client which serves the purpose of identifying customers who can potentially default in device loan payment. Data provided from the client for this project was poorly structured, and inconsistent. A majority of time was invested in cleaning the data and extracting useful features by deep feature engineering. PCA was done to reduce noise in the data, experimentation with different sampling ratios was conducted for dealing with imbalance in the data. After a lot of experimentation with different algorithms, we settled with Random Forest, which had optimal bias-variance trade-off. Precision achieved for the same was more than 75%. The behavioral data used for the model was less than 6 months old.

### **Amazon Development Centre, Hyderabad – Risk Analyst**

JULY 2016 – FEBRUARY 2018

#### **Job Description**

- Identifying, tracking, and reporting instances of non-compliant consumer processing.
- Monitoring transactional activity in order to identify potentially fraudulent and high risk activities conducted by customer entities.
- Conducting investigations in order to assess and mitigate risk of monetary loss stemming from processing activities.

## **PERSONAL LEARNING PROJECTS**

### **Machine Learning**

- Time series analysis with mobile game data. (Rolling window estimations, Exponential smoothing, Time-series cross validation, parameters selection, Getting rid of non-stationarity, SARIMA implementation, Linear models, Feature extraction, Regularization, feature selection, XGBoost)
- Stock Price Prediction with LSTM and Prophet.
- Recommender systems on Movielens dataset and Restaurant- customer Dataset from UCI ML (Content based filtering, Collaborative filtering, PCA, Matrix Factorization for feature engineering, Word vectors, similarity based algorithm, clustering, hyper parameter tuning).

### **Computer Vision**

- Image Classifier with convolution neural networks.
- Similar Image Search System with CNN and Autoencoders. Associated this system with apparel recommendation engine to get similar items.
- Facial Recognition System with Facenet, ArcFace and KNN Classifier.
- Blood Cell Detection in microscopic images with Faster RCNN.

## Natural Language Processing

- Sentiment Analysis on Amazon Food Reviews Dataset. (Text Processing- Stemming, Stop-word removal, Tokenization, Lemmatization, Bag of Words, n-gram, tf-idf, word2vec.)
- Named Entity Recognition with Char-bi-LSTM
- Text Classification with TextCNN. Sequence-to-sequence modelling for machine translation.

## SKILLS

- **Machine Learning:** GLM, KNN, SVM, Clustering, Ensemble models (GBM, XGB, RF).
- **Deep Learning:** CNN, LSTM, Transfer Learning
- Apache Spark
- **OS:** Redhat, Ubuntu, Windows
- **Languages:** Python
- **Databases:** MySQL, PostgreSQL
- **Cloud Services:** AWS, GCP
- **Libraries and Frameworks:** Pandas, numpy, scikit learn, scipy, statsmodel, Turicreate, OpenCV, dlib, NLTK, gensim, prophet, tensorflow, keras, h2o, MXNet, Django (expertise: Competent)

## EDUCATION

### Anna University, Chennai **M.Tech, Biotechnology**

JUNE 2014 - JUNE 2016

CGPA: 8.66/10

#### Relevant Courses:

- Computational Biology
- Bio-statistics

#### Thesis:

Study on *Eis* protein, a virulence factor from *Mycobacterium tuberculosis*, using computational approaches.

### Heritage Institute of Technology, WBUT, Kolkata **B.Tech, Biotechnology**

JUNE 2010 - MAY 2014

CGPA: 7.45/10

#### Relevant Courses:

- Database Management
- System
- Programming with C
- Engineering Mathematics
- Numerical Analysis
- Bioinformatics

#### Thesis:

Action of Dextranase extracted from *Tinospora cordifolia* on *Dental Plaque* and its formulation in Toothpaste.