Somak Dutta

+91 - 9820597793
somak.dutta@outlook.com
www.linkedin.com/in/somak-dutta

- **11 years'** experience in **Java/ J2EE / Python** applications across Insurance, Retail and Investment Banking domain
- Previously worked as **Senior Analyst, Bank of America Continuum India, Mumbai**
- **Top 60% in Kaggle** competitions and **top 4% in Hacker Earth** Machine Learning challenges
- Completed **IBM Data Science Specialization** in Coursera
- Worked with Supervised/Unsupervised learning, Natural Language Processing (NLP) and Image Classification using deep learning networks (CNN)
- Git: *https://github.com/in-tandem/*   Kaggle: *https://www.kaggle.com/somaktukai*

Skill set

- Languages/Framework: Java (jdk 8), J2EE, EJB, JMS, JAX-WS (Soap webservices), Jersey Rest, Hibernate, XML, JSP/Servlets, JDBC, PL/SQL, OSB, Python, Flask, Numpy, jQuery, Kendo UI, angular 4,GSON
- Machine Learning Toolkits: Pandas, matplotlib, scikit-learn, xgboost, lightgbm, imblearn, Keras, Folium, NLTK, Gensim, Topic Modelling, NLP, Deep Learning, Artificial Neural Networks, Convolutional Neural Network, Anaconda
- Unit Testing Tools Used: Junit , Jmock, Eclemma, Selenium
- Code Analyzer: PMD, Visual VM, Fortify
- Build Tools / Servers: Ant, Maven, Team City, Jenkins
- Database: Oracle 10g /11g, Netezza, Oracle Exadata
- Application Server: Weblogic 9.2 / 10.3 / 11g, Tomcat 8
- IDE: Eclipse, JDeveloper, Visual Studio Code, Jupyter Notebooks, Spyder
- Repositories: GIT, SVN, VSS
- Methodology & Agile Tools: Agile, Waterfall Model, JIRA, Version One

Work Experience

- Senior Analyst, Bank of America Continuum India (20th Oct2014 till 7th September 2018)
- ENO at Credit Suisse, Mumbai. (24th Feb 2014 – 17th Oct 2014)
- Senior Consultant at Oracle Financial Services Software Ltd, Mumbai. (30th Aug 2010 – 5th Feb 2014)
- Assistant Systems Engineer at Tata Consultancy Service, Kolkata. (13th Aug 2007 – 13th Aug 2010)

Educational Qualification

- B Tech – Computer Science and Engineering. From West Bengal University of Technology, 2007
- ISC – CISCE, New Delhi, 2003 & ISCE – CISCE, New Delhi, 2001

Personal Details

- Sex: Male
- Date of Birth: 27th September 1984
- Current Location : Kolkata India

Assignment Summary

Kaggle / Hacker Earth and Other Capstone Projects	
Kaggle Do Not Overfit Challenge ✪ Goal: Classification challenge with a goal to reduce overfitting ✪ Classification Type: Binary classification ✪ Public leader board - top 67%	- Final Submission Score: 0.724 Metric used : AUC ROC Score - Model used: Logistic Regression classifier - Cross validation used: stratified k fold with 5 fold - Parameter tuning approach: gridsearch - Pearsons correlation coefficient was used to determine top 50 features to include in final training set - Majority classes were down sampled using imblearn library - 5 fold stratified cross validation was used to further reduce impact of overfitting
Kaggle ELO Merchant Category Regression Challenge ✪ Goal: Prediction of customer loyalty ✪ Regression ✪ Public leaderboard - top 62%	- Final Submission Score: 3.63 Metric used : RMSE - Model used: LGBM Regressor - Cross validation used: k fold with 10 fold - Parameter tuning approach: gridsearch
Societe General Brainwave Hackerearth Machine Learning Challenge ✪ Goal: Classification of Complaint Handling for the given data set ✪ Classification Type: Multi class Text Data ✪ Public leaderboard - top 4%	- Final Submission Score : 0.72 ▪ Metric used : F1 Weighted - Model used: DecisionTreeClassifier with depth of 8 and criteria as GINI - Cross validation used: stratified k fold with ten fold - Parameter tuning approach: gridsearch - NLTK and Gensim libraries were used to perform topic modelling as a technique to feed language text to model
IBM Capstone: Battle of the Neighbourhood	- Model used: KMeans Clustering - Toronto Culture and Community Center data set was downloaded from government website

✪ Take any open source data set available and perform analysis and report out the findings ✪ Clustering	▪ Using geocoder to get latitude/longitude ▪ Proper data analysis was performed with results and conclusions linked up to hypothesis

Saber2 Application Development , Senior Analyst – Bank of America Continuum India

Languages:	
Java, Jersey Rest API, Python, Flask API, Jinja Templates, jQuery, kendo UI, angular 4, bootstrap, JDBC Database: Netezza/ Oracle Exadata /Redis/ in house non-relational DB Sandra	Design, development and production support of Web Submission Framework Enabling multiple work streams (10+) users to manually upload their data, undergoing configurable validations, business rules for each template
	Design, development and production support for Simplify and Improve project for Enterprise Capital Liquidity Management team. A configurable rule engine was built for the users to enable mapping of unstructured data to a structured format for downstream consumption
	Development of various functionalities for Control Panel system for execution and monitoring of Basel 3 calculators and supporting functions.

Trade Management IT , ENO – Credit Suisse

Languages:	
Java, JMX, EJB Pollers, JAX-WS, Ant, Python, Team City, Hibernate, JDBC Database:Oracle	Design and development of JMX Monitoring tool to expose WebLogic application services monitoring capabilities using Mbeans
	Design and development of automated regression tool. Tool would create a system integration environment, creation and fulfilment of WebLogic domain, creation of mini schema and execution of test cases and various other integration tests
	Integration with clearing houses in US and EMEA regions for Credit Default Swap trades

Sun Corp Project, Senior Consultant - Oracle Financial Services Software Ltd

Languages:	
	Integration of Banks Document Management Systems with Term Deposits Module Customization to fulfill requirement of Single Account Single Deposit for TD products

Java, JAX-WS, OSB, Hibernate, Ant, Hudson Database:Oracle	Development of Credit assessment for Loan Application Integration with third party system to calculate capacity to repay for Loan applicants Integration with third party system to fetch credit bureau records for Loan applicants

National Australian Bank Project, Senior Consultant - Oracle Financial Services Software Ltd	
Languages: Java, JAX-WS, OSB, Hibernate, Ant, Hudson, ADF Database:Oracle	Integration with CRM Service Requests Module for Customer Origination
	Integration with Oracle Customer Hub for user management and provisioning
	Creation of various Loan calculator simulators and UI screens

TCS Bancs Insurance , Assistant Systems Engineer - Tata Consultancy Services	
Languages: Java, PL/SQL, JDBC Database:Oracle	Development of member upload feature, coinsurance module, upgrades to accounting module, batch development such as underwriting batch, reporting batches.

Recognition

- Multiple gold/silver awards as recognition for work done in WSF/SIM project
- Recognition from Management for work done in Tech Hygiene and JMX Tool
- i-appreciate award for recognition of work done in Oracle Financial Services
- On the Spot award for recognition of work done in TCS

Certification

- Coursera: IBM Data Science Professional Specialization