

# ISHAN PRADIP BORKER

## Data Scientist

[ishan123ppp@gmail.com](mailto:ishan123ppp@gmail.com) | [+91 8600313943](tel:+918600313943) | [in LinkedIn Profile](#) | [Github Profile](#)

### Profile

- IT professional with 3 years of experience. Seeking to advance my career in Data Science, Machine Learning and effectively apply my professional, technical skills in diverse portfolio
- Keen to learn new technologies and take up challenging roles

### Technical Skills

**ML libraries:** Numpy, Pandas, Tensorflow, Keras, SciPy, seaborn, Scikit-learn, Plotly-dash, Matplotlib

**ML algorithms:** Classification, Regression, Clustering, Boosting, Time series forecasting

**Programming languages:** Experience in Python, Well verse with R, C++, Java, JSON Parsing

**Operating Systems:** Linux (Ubuntu, CentOS), Windows, macOS

**Technologies:** HTML, JS, PHP, Jupyter notebook, Google colab, FHIR, POSTMAN, Github, Gitlab, Docker, Homomorphic encryption

**Database Technologies/BI tools:** MySQL, SpagoBI, Tableau, Power-BI, GNU-plot, Excel

**Skills:** Flexible, Eager to learn new technologies, team player

**Hobbies:** Playing indoor, outdoor games, listening to music, reading books, newspapers, magazines, watching the news, movies

**Languages Known:** English, Japanese (N5 level), Hindi, Konkani, Marathi, Sanskrit

### Work Experience (3 years)

**Molecular Connections Pvt. Ltd., Bengaluru, Karnataka** *Senior Software Engineer – Data Science (Full-time)* 07/2020-10/2020

**EHR Predictive Modelling:** Performed cleaning EHR Patients data, EDA, created visualization and correlation on varied features

- Generated lab value fields based on specific patient IDs, Lab names, performed binary encoding concerning lab values
- Trained the machine learning model to calculate accuracy, precision, exported the model on joblib/pickle, explored plotly-dash

**FHIR HL 7:** Implementation of modules like Formal Concept Analysis, Concepts, parsing of JSON files using python, FHIR Parser

- Worked on HAPI FHIR Server, creating resources from JSON, testing the REST operations on POSTMAN
- Created own FHIR Server, tested resources in POSTMAN, executed python script to upload JSON files, created bundle transactions

**Rubixe, Bengaluru, Karnataka**

*Data Science Consultant*

12/2019-06/2020

**Employee Performance Analysis:** Cleaned the data, performed EDA, created a visualization to find department wise employee performance, factors affecting it, built models to improve it using machine learning algorithms like Random Forest, Gradient Boosting, XGBoost, ANN, KNN, Logistic Regression, SVM, Decision Tree with highest accuracy, precision of 96% using Gradient Boosting, used feature engineering, Grid & Randomized Search Cross Validation(CV), SMOTE to improve model performance

**Spare Parts Inventory Management:** Cleaned, analyzed the spare parts data, created visualization using python libraries, built the model using time series forecasting algorithms, achieved better predictions for ARIMA model

**Improving IT Service Management:** Imported data from server using SQL queries, performed EDA, implemented prediction on priority tickets and reassigned tickets using machine learning algorithms with highest accuracy of 86% and 84% respectively

- Implemented algorithms like ARIMA and SARIMA on the incidents, achieved better predictions using Rolling Forecast

**Sales Prediction:** Used techniques like SMOTE for handling imbalanced datasets, PCA for dimensionality reduction

- Predicted the status of sales potential using Random Forest, XGBoost, SVM, ANN, KNN, Logistic Regression, Gradient Boosting, Decision Tree, Extra Trees, Naïve Bayes algorithms, achieved 72% accuracy, plotted ROC, precision-recall curve

**Telecom Churn Prediction:** Predicted churn using Random-Forest, XGBoost, Gradient Boosting, Decision Tree, Extra Trees, Logistic Regression, achieved the highest accuracy of 97%, 95% precision, used K-fold, Stratified K-Folds CV to improve performance, plotted ROC, precision-recall curve, calculated churn-risk score using probability factor, exported model using joblib

**Bank Good Credit Score:** Imported Customer Accounts, demographics, enquiry data from server using MySQL, created visualization to understand what influences the customers having good or bad credit history, implemented Random Forest, Decision Tree, Extra Trees, XGBoost, Gradient Boosting, Logistic Regression, SVM, ANN, achieved 95% accuracy, 91% precision, merged Customer Accounts with Customer Demographics, achieved highest 99% accuracy, precision, also merged Customer Enquiry with Customer Demographics, achieved highest 99% accuracy, precision

**Digiapt Software Technologies Pvt. Ltd., Bengaluru, Karnataka** *Associate Software Developer (Full-time)*

10/2019-10/2019

- **Gathering and filtering of companies' data:** Used Google spreadsheet to gather company's data from MCA, Tofler etc., worked on Tableau to filter them, used My maps to plot company's location on Google maps, used Scrapy to scrap websites
- **Capture Email:** Snovio to capture email of employee from company using csv file, achieved employee's official email with the designation

**Eaglys Inc., Yoyogi, Shibuya, Tokyo, Japan** *Research Engineer (Full-time)* 12/2018 - 07/2019

- **Back Office system for Man-hour Management:** Captured attendance spreadsheet in Google Drive using python, converted into csv, performed summation, generated graphs, PowerPoint presentation
- **SecureDB Product:** Tested SQL queries for real estate, successfully implemented fuzzy keyword searching on the encrypted database using python, C++, performed API documentation for SecureDB using Natural Docs and PyDoc

**Teczuno Global India Pvt. Ltd., Bengaluru, Karnataka** *Data Analyst (Full-time)* 04/2018 - 08/2018

- Gathered data from mygov.in, helped in application development to locate start-up company using Tableau, built concepts using the data from sectors like Healthcare, Automotive, Agriculture as part of pre-sales support

**Anant Infomedia Pvt. Ltd., Panaji, Goa** *Project Trainee (Full-time)* 08/2017- 02/2018

- **Predicting attrition of employees using Microsoft Azure Machine Learning:** Designed models using machine learning algorithms like classification, regression, clustering, anomaly detection for different pharmaceutical clients
- **Data Analysis using SpagoBI:** Used MySQL, SpagoBI for Data Analysis, created dashboards, cockpits for compliances, developed OLAP, BIRT, deployed them on the SpagoBI server, designed codes using PHP, JS, HTML, views using MySQL on Navicat

### Internship Experience

**Rubixe, Bengaluru, Karnataka** *Data Science Intern* 03/2020 - 05/2020

- **AI for Hiring:** Designed business proposal for student's academics & placement data, built model using classification algorithms to predict students placed, published a research paper

**Goa Electronics Limited, Panaji, Goa** *Project Intern* 10/2018 - 11/2018

- Analyzed compliant government websites, designed SRS for labour & employment, land dispute, fire & emergency services, created flowcharts for employment exchange services

**C-DAC ACTS, Pune, Maharashtra** *Project Intern* 07/2016 - 05/2017

- **Parallelization of Backpropagation algorithm using OpenMP:** Achieved better efficiency, performance of parallel programs on Intel Xeon Phi Processor, carried out the dissertation work, published a research paper

**CSIR- National Institute of Oceanography, Goa** *Project Intern* 12/2013 - 01/2014

- Research on robotic OS, worked on serial port C program to transfer data from one USB port to another

### Educational Background

- **2015-2017: Master of Technology (High Performance Computing),** C-DAC Pune, Veltech University Chennai - **8.61 C.G.P.A.**
- **2011-2015: Bachelor of Engineering (Computer Engineering),** Goa College of Engineering - **70%**
- **2010-2011: Higher Secondary School (Class XII),** Goa Board - **74%**
- **2008-2009: High School (Class X),** Goa Board - **82.17%**

### Workshops/Certifications

- **Deep Learning specialization** by Coursera and deeplearning.ai on coursera.org
- **Certified Data Scientist & Data Science Foundation** by International Association of Business Analytics Certification, Amsterdam
- **Certified Data Scientist** course completion certificate by DataMites, Bengaluru
- Presentation on "**Backpropagation Algorithm & use of OpenMP in ML**" in International Workshop on IOT & TV White Spaces
- 1<sup>st</sup> prize in National Summer Training Program on "**Bigdata & Hadoop**" by Revert Technology Pvt. Ltd., EDC-IIT Roorkee
- Diploma Course in Java programming with JDBC, C programming at MICE
- Research paper publication "**Parallelization of Backpropagation algorithm & Benchmarking**" in IJCR, June 2017
- Research paper publication "**Artificial Intelligence for Hiring**" in International Journal of Science & Research, Nov 2020
- AMCAT examination - 77.72 percentile

*References and more information will be provided on request*