# ISHAN P. BORKER

✉ **ishan123ppp@gmail.com** | ☎ **8600313943** | in **LinkedIn** | ○ **GitHub**

## Key Skills

### Generative AI

RAG
LangChain
LLM
AWS Chatbot
Amazon Bedrock
BrowserBase
Perplexity.AI
Mosaic AI

### Machine Learning

Python
FastAPI
Data Science
Supervised Learning
Unsupervised Learning
Computer Vision
NLP
Linear and Logistic
Regression
Clustering algorithms
Decision Trees
Neural Networks
Reinforcement Learning
Feature engineering
Statistical analysis

### Machine Learning libraries:

Tensorflow
Keras
SpaCy
PyPDF
PyLDAVis
BERTopic
NLTK
PySpark
MLFlow

### Algorithms

Classification
Regression
Clustering
Time series forecasting
Yolov3
Yolov4
OpenCV
Mask R-CNN
Faster R-CNN
Named Entity
Recognition

## PROFILE

- Total experience of more than 7.5 years in software development of various platforms like Python, data science, machine learning, Generative AI, natural language processing (NLP), research and development, C++, FastAPI, MySQL, Tableau, Network Service Orchestration. Part of the role includes analyzing requirements, architecting the entire application, providing end-to-end solutions to customers from solution architecture, developing, implementing the application, and reviewing using technologies like POSTMAN, Docker, Linux/Unix, JSON scripting, Golang and design to deployment
- 6.5 + years of ML experience plus 7.5 + years of IT experience
- Experience studying client requirements and recommending algorithms to be used. In-depth experience handling feature engineering and creating visualizations
- Experience in end-to-end delivery of project from the proposal stage, requirements gathering, assisting in design, implementation, technical documentation, excellent understanding of customer engagement and delivery
- Highly adaptable and dedicated to solving business problems, Flexible, team player, keen to implement new technologies as per customer needs, ready to take up challenging roles
- Seeking to enhance my career in AI and effectively apply my technical and management skills in diverse portfolios

## EXPERIENCE

**Persistent Systems, Verna, Goa as Lead Software Engineer, May 2022 - Present**

**Role: Apr 2024 – Present** as **Lead Software Engineer**

- **Generative AI Platform development - Co engineering – Data (Client: Baldwin Risk Partners)**: <u>Classification</u>: Running the RAG classification and extraction code for different policy documents. Exploring the chunking strategy for the large tables obtained after extracting the documents using LLM calls
  <u>Integration of Amazon Bedrock Agent and AWS Chatbot with Microsoft Teams</u>: Changing the email while testing the message in Amazon SNS and publishing it inside Teams Channel, creating a knowledge base and Amazon Bedrock Agents and testing it with prompts, creating an application inside the developer portal in Microsoft Teams for integration
  <u>Actionable Web Technologies</u>: Creating a tabular analysis for evaluating different actionable web technology tools like BrowserBase, Maxun based on future scalability, and browser dependencies. Python code to enter the URL and download the relevant PDFs, code to connect the LLM model to Amazon Bedrock, and then using BrowserBase packages like Playwright and its APIs to send the URL to download the relevant PDFs**.** Working on stealth mode by modifying the existing code to bypass the captcha or login methods and capture the link, file name, and other metadata. Running the same code on the top 5 carriers being implemented under Carrier Document Download, downloading policy document for each agency code using credentials for CNA and selective flood carrier website and integrating metadata to it, performing download policy document using LLM prompts and Stagehand approach on Hartford carrier website
  <u>Capabilities of Perplexity.AI</u>: Working on improving prompt caching, dynamic updates, response accuracy, cache persistence and query handling using perplexity.ai, retrieving the response format obtained from perplexity API, and checking for the separation of the citations from the text. Testing feasibility of Perplexity in BRP Copilot UAT and potential business benefits
  <u>Mosaic AI Model Fine-Tuning</u>: Performing model fine-tuning on the RAG Component Classification and Extraction dataset from S3 bucket using Databricks Notebook, performing model testing using gpt-35-turbo-SLM, training the fine-tuning run using meta-llama models, deploying the fine-tuned model to serve the endpoint and testing it
  Worked on the client requirements like risk map automations, combination of structured and unstructured sources, graph DB for automating the RFP process, LLMs with structured knowledge

Relationship Extraction
Topic Modelling
Latent Dirichlet
Allocation
Fasttext

**Other Skills**

Actionable Web
Technologies
JSON
HTML
JS
PHP
Curl commands
Unix/Linux commands
Golang

**OS**

Linux/Unix
Windows
macOS

**Database/BI tools**

MySQL
SpagoBI
Tableau
PowerBI
Navicat
GNU-Plot

**Python Libraries**

NumPy
Pandas
SciPy
Seaborn
Scikit-learn
Plotly-dash
Matplotlib

**Tools/Environments**:

EMS
Crosswork Workflow
Manager (CWM)
Network Service
Orchestration (NSO)
Lux
Jupyter Notebook
Google Colaboratory
Kaggle Notebook
Databricks Notebook
VSCode
PyCharm
Spyder
GitHub

graphs, graph RAG, knowledge augmented graph and preparing the documentation. Prepared use case presentations for all the POCs.

- **CISCO NSO Engineering (Client: CISCO):** Worked on EMS adapter activity support for different workflows like distribute, activate, commit image on IOS-XE device, working on API hardening: adding error handling conditions to the EMS adapter Go code, validating the inputs for SWIM APIs, performed code refactoring on it, performed add image, activate and commit using install method for IOS-XE device, updating the EMS adapter to be compliant with CW infrastructure by refactoring the code using Authentication, No Authentication, running different EMS APIs inside cw-admin SVM Infrastructure like import, activate, distribute and commit image using curl commands and updating the wiki page, worked on the unit test cases for the refactored codes, developed POC for Go client for fetching the CCO catalog and SMU details for each record in catalog, fetched multiple devices inside the jobsummary EMS Code

**Role: May 2022 – Mar 2024** as **Senior Software Engineer**

- **Routermation CISCO NSO Offshore Transition (Client: CISCO)**: Worked on Crosswork Workflow Manager (CWM) installation, creation of adapters, resources, secrets, EPNM, created workflows, adapters for different applications and tested it on POSTMAN, Captured requests and responses for different APIs like import, distribute, activate and commit image, get job summary, get device details, device upgrade, consolidated workflow etc. for Element Management Systems (EMS) on POSTMAN, implemented codes for networking device upgrade and created adapters using Golang and executed the workflows in CWM UI, executing the SWIM APIs in CWM UI instance, worked on Network Service Orchestration (NSO) installation and Routermation Automation Kit for device onboarding, fleet upgrade and golden config, performed LNT device upgrade using RAKfu Main workflow and lux test cases, fixed critical and major issues for CW-EMS adapter using SonarQube, updated the CW-EMS adapter with CWM 1.1 support and library integration, worked on IOS-XE device upgrade

- **PDF Text Extraction and Code Testing (Client: IBM Infosphere)**: Worked on PDFs from URLs. Extracted, and preprocessed the text from it using PyPDF and generated histograms using nltk. Worked on extractive text summarization using different summarizers like LSA, Luhn, Text Rank, and Lex Rank and calculated the inference time of each summarizer. Tested FastAPI codes for training and inference concerning Fasttext classifier on text classification dataset with categories like Chemical Hazardous, Restricted Prohibited, Neither, calculated the total training time of all the models. Performed Volumetric Stats for Catalog files like unique UNSPSC for Primary, Corrected Category, number of data points available and missing for each UNSPSC for corrected category

- **Serving NER and REL models using SpaCy**: Worked on clinical trials dataset using SpaCy for Named Entity Recognition and Relationship Extraction using pipelines like tok2vec, different transformers, plotted loss curves, improved performance metrics for all models. Served English SpaCy model on FASTAPI and tested it on POSTMAN and Swagger UI. Served own trained NER, REL models using FASTAPI, tried them on POSTMAN, deployed them on AWS EC2, dockerized FASTAPI NER code, pushed docker container on GitHub

- **Topic Modelling**: Worked on topic extraction on docoh dataset of companies (Intuit, UHG, Wells Fargo, etc.) using packages like BERTopic, Latent Dirichlet Allocation (LDA), and Contextual Topic Modelling (CTM). Performed dimensionality reduction (PCA, UMAP), and clustering (HDBSCAN) techniques, using transformers like Roberta, finbert, and sentence transformers to improve topic extraction and visualizing the topics using PyLDAvis. Development is done using Python inside Google Colaboratory

- **Binary classification of network attacks**: Worked on binary classification of network attacks present in NSL-KDD dataset, performed EDA, data pre-processing for different features using matplotlib, seaborn, found statistical insights, performed feature engineering using one-hot encoding, label encoding, feature scaling, normalization, performed feature selection using methods like filter, wrapper and embedded, performed model selection using cross-validation, bootstrapping, used different machine learning classifiers like Random Forest, Decision Tree, XGBoost, Gradient Boosting, Voting Classifier and trained the model, performed hyper-parameter

Amazon EC2
Azure ML Studio
POSTMAN
Swagger UI
GitHub Copilot
AWS Code Whisperer
SonarQube
Docker

**Languages**
English
Hindi
Konkani
Marathi
Sanskrit
Japanese (Basic)

**Date of Birth**
11/08/1993

**Address**
"ABHA", H.N.1794,
Walkeshwar Wada,
Betim, Bardez, Goa -
403101

---

tuning using GridSearchCV and RandomizedSearchCV, evaluated performance metrics, saved the trained model and exported it on the server using joblib, worked on model explainability using eli5, Partial Dependence Plot, Individual Conditional Expectations, Local Interpretation of Model Explanation, SHAP

- **Smart Interviewer**: Developed MCQ part of the test using Python, created a dataset of question bank using .toml for AI Technothon event. Development is done using Python, GitHub Copilot, and VS Code

- **Image Segmentation**: Worked on Medical Image segmentation using U-NET, Mask R-CNN on NVIDIA GPU Cloud (NGC)

**TechnoPro India, Bengaluru as Machine Learning Engineer, Dec 2020 - Apr 2022**

- **Annotation (Client: Nexar)**: Worked on semi-automated annotation using CVAT, ML algorithms like YOLOv3, Faster R-CNN, Mask R-CNN, and manual annotation using tools like Super Annotate, VoTT, LabelBox, etc. Led a team of annotators, and trained them on annotations like bounding box, polygon, polyline, cuboid, video, 3D annotation, instance, and semantic segmentation

- **Smart Farming**: Implemented tomato image classification using SVM, KNN, MLP, LR. Development is done using Python inside the Jupyter lab

- **Industry 4.0 (Client: Toyota Kirloskar Motors)**: Worked on detecting car parts in images and trained model using YOLOv3, tested code on car damaged parts detection using detectron2, Mask R-CNN, worked on predictive maintenance of bearing using regressors like SVR, RF, DT. Worked on Object Detection and Tracking of real-time intrusion using YOLOv4, OpenCV. Development is done using Python inside Google Colaboratory

- **Perimeter Intrusion Detection System (Client: Bengaluru International Airport Limited)**: Worked on POC, designing RFP on perimeter intrusion detection project for Airport surveillance covering aspects of Machine learning, Computer vision, RADAR for video surveillance, verification, tracking, access control security, and Smart Command and control system

- **Reviewer Analysis (Client: Nexar)**: Analyzed reviewer incident data in graphs using Python, matplotlib, and Seaborn, designed presentations, case studies, RFP, and POC for different AI-ML projects

**Molecular Connections, Bengaluru as Senior Software Engineer, Jul 2020 – Oct 2020**

- **EHR Predictive Modelling**: Cleaned EHR data, performed EDA, visualization, correlation, binary encoding on lab values, trained model, and exported model on the server using joblib, worked on FHIR Parser using Python, built HAPI FHIR Server by creating resources from JSON, tested REST operations on POSTMAN, created own FHIR Server

**Rubixe, Bengaluru as Data Science Consultant, Dec 2019 - Jun 2020**

- **Employee Performance Analysis**: Cleaned the data, performed EDA, created a visualization to find the department-wise performance of employees, and built a model to improve performance using ML algorithms like Random Forest, Gradient Boosting, XGBoost, ANN, KNN, LR, SVM, DT, used feature engineering, Grid & Randomized Search CV, SMOTE

- **Spare Parts Inventory Management**: Cleaned, analyzed the data, created the visualization, built a model using time series algorithms, and achieved better predictions for ARIMA

- **Improving ITSM**: Imported data from the server, performed EDA, implemented prediction on priority tickets, reassigned them using ML algorithms, used ARIMA, and SARIMA on incidents, achieved better predictions using Rolling Forecast

- **Sales Prediction**: Used SMOTE to handle imbalanced dataset, PCA for dimensionality reduction, predicted sales status using ML classification algorithms, plotted ROC, and precision-recall curve

- **Telecom Churn Prediction**: Predicted churn using ML classification algorithms, used K-fold, Stratified K-Folds CV to improve performance, plotted ROC, precision-recall curve, calculated churn-risk score, exported model on the server

- **Bank Credit Score**: Imported Customer Accounts, Demographics, and Enquiry data from the server, created a visualization to check factors influencing customers having a good or bad credit history, achieved 95% accuracy using ML classification algorithms

**Digiapt Software, Bengaluru as Associate Software Developer, Oct 2019 - Nov 2019**

- Categorized the data of companies using Tableau, plotting their locations using My Maps, used Snovio to capture emails of employees and analyze roles

**Eaglys Inc. (Deputed from Teczuno), Tokyo, Japan as Research Engineer, Dec 2018 – Jul 2019**

- **Back Office Manhour Management**: Generated reports using graphs, and presentations for attendance spreadsheet from Google Drive

- **Secure DB**: Tested SQL queries, achieved API documentation using Natural Docs, PyDoc

**Teczuno Global India, Bengaluru as Data Analyst, Apr 2018 - Aug 2018**
- Gathered data, developed an application to locate startups using Tableau, and built-up concepts as a part of pre-sales support

**Anant Infomedia, Panaji, Goa as Project Trainee, Aug 2017 - Feb 2018**

- **Employee Attrition**: Predicting employee attrition using classification, regression, clustering, and anomaly detection inside Azure ML Studio

- **Business Intelligence**: Used SpagoBI for data analysis, created dashboards, cockpits, reports and deployed on SpagoBI server, worked on PHP, JS, HTML, Navicat, MYSQL

**INTERNSHIP EXPERIENCE**

**Rubixe, Bengaluru, Data Science Intern, Mar 2020 - May 2020**
- **AI for Hiring**: Designed business proposal for student placement, built model using machine learning classification algorithms for placement prediction, and published paper in International Journal of Science and Research (IJSR)

**Goa Electronics Limited, Goa, Project Intern, Oct 2018 – Nov 2018**
- Designed Software Requirement Specification (SRS) for labor-employment, land dispute, and fire-emergency services, and created flowcharts for employment exchange, worked on the analysis of making compliant government websites

**C-DAC ACTS, Pune, Project Intern, Jun 2016 – May 2017**
- MTech Project Work: Implemented code of backpropagation using C in Ubuntu, used OpenMP technique to execute parallel code to improve efficiency, benchmarking done on Intel 64-bit, Intel Xeon architectures, achieved better performance on Intel Xeon Phi, carried out dissertation work and presented a paper in the International Workshop on Internet of Things and TV White Spaces (WIOT' 2017), published research paper in the International Journal of Current Research (IJCR)

**CSIR-NIO, Goa, Project Intern, Dec 2013 – Jan 2014**
- Research on robotic OS, implemented a serial port program to transfer data from one USB port to another in Ubuntu

## EDUCATION

- M. Tech High Performance Computing at C-DAC and Veltech University, 2015 – 2017, 8.61 CGPA
- B.E. Computer Engineering at Goa College of Engineering (Goa University), 2011 –2015, 70%

## CERTIFICATIONS

- Introduction to Large Language Models (LLMs) In Python by Udemy
- MCP Crash Course Complete Model Context Protocol in a Day by Udemy
- Intro to Large Language Models (LLMs) by Udemy
- Participated in Semicolons 2025, Annual Global Hackathon of Persistent Systems
- Python ML Competency Level - 3 Certification by Persistent University
- Python ML Competency Level - 2 Certification by Persistent University
- Automate the Boring Stuff with Python Programming by Udemy
- Complete Python Bootcamp 2024: Zero to Expert in Python by Udemy
- AWS SageMaker Practical for Beginners Build 6 Projects by Udemy
- MongoDB – The Complete Developer's Guide 2024 by Udemy
- Complete Tensorflow2 and Keras Deep Learning Bootcamp by Udemy
- Taming Big Data with Apache Spark and Python – Hands-on by Udemy
- Natural Language Processing with Python by Udemy
- Natural Language Processing with Transformers in Python by Udemy
- LangChain with Python Bootcamp by Udemy
- REST APIs with Flask and Python in 2024 by Udemy
- Python for Data Science and Machine Learning Bootcamp by Udemy
- Complete Data Science, Machine Learning, DL, NLP Bootcamp by Udemy
- Git for Geeks: Quick Git Training for Developers by Udemy
- Amazon Bedrock & AWS Generative AI [Beginner to Advanced] by Udemy
- Generative AI: Beginner to Pro with OpenAI & Azure OpenAI by Udemy
- LangChain- Develop LLM powered applications with LangChain by Udemy
- Generative AI for Practitioner Certification Course (Objective and Subjective) by Persistent University
- Generative AI Foundation – Generative AI Assisted Coding using Amazon Code Whisperer by Persistent University
- Databricks Accredited Generative AI Fundamentals
- Career Essentials in Generative AI by Microsoft and LinkedIn
- Databricks Lakehouse Machine Learning Associate
- Participated in Semicolons 2023, Annual Global Hackathon of Persistent Systems
- Fundamentals of the Databricks Lakehouse Platform Accreditation
- Deployment of Machine Learning Models by Udemy
- Paper publication "Artificial Intelligence for Hiring" in IJSR
- Certified Data Scientist & Data Science Foundation by IABAC, Amsterdam
- Deep Learning Specialization by Coursera and deeplearning.ai on Coursera.org
- Certified Data Scientist course completion certificate by DataMites, Bengaluru
- Paper publication "Parallelization of backpropagation algorithm & Benchmarking" in IJCR
- Paper presentation on "Backpropagation Algorithm & use of OpenMP in ML" in International Workshop on IOT & TV White Spaces