

Vinit Ganorkar

AI Lead | Data Scientist | ML Engineer

✉ vinit.ganorkar@gmail.com ☎ +919028381267 📍 Bangalore, India 📁 Portfolio [in](#) LinkedIn [@](#) Git

Professional Summary

- Almost **8 years of experience** in working on variety of machine learning (ML), computer vision (CV) and Natural language Processing (NLP) projects within Insurance, Investment, Telecom and Medical domain.
- Ability to devise a structure for **solving Unstructured Business Problem**.
- Hands on knowledge of working with **Big data of petabyte scale using distributed systems ~ 400M users**
- Substantial experience of **working with stakeholders** to identify new requirements and **opportunities**.
- **Lead teams** to develop, adapt, combine, and choose data science methodology (statistics, machine learning, optimization, simulation) to address complex modeling requirements.
- **Highly self motivated individual** - currently learning more about - GenAI, GNN, ChatGPT, EdgeAI etc

Accomplishments

- Built a customer attrition model for **Jio Mobility** business which helped in avoiding potential revenue loss of **60 Million INR per quarter** from Mumbai region alone.
- Developed **Jio phone** churn model which extensively helped device team in understanding the root cause of churn, resulted in **increase in customer retention rate** by a great extent.
- Churn Prediction model for **Jio Fiber** business helped in avoiding revenue loss of **40 Million INR per month** also it helped customer care team to do probability based targeted out bond calling.
- Developed a text classification model to classify user feedback into various themes resulted in **increased efficiency** of **NPS** team.

Achievements

Jul 2022	Winner of Hewlett Packard Enterprise ML Hackathon - ML 🔗 <i>Hewlett Packard Enterprise</i>
Aug 2019	Runner up of Colombia Gender Recognition Challenge - NLP 🔗 <i>Times Internet</i>
Nov 2018	Finalist of Microsoft AI Challenge (Ranked 32) - Information Retrieval 🔗 <i>Microsoft India</i>
Sep 2018	Winner of Springer Nature Pune Artificial Intelligence Hackathon - CV 🔗 <i>Springer Nature, Pune</i>
Jul 2018	Winner of Data Science Edition 2 Hackathon - Sequence Prediction 🔗 <i>Techgig</i>
Dec 2017	Winner of News Article Recommender Challenge - Click Prediction 🔗 <i>IIT Powai - International AI Summit 2017</i>
Jun 2017	Runner up of Big Data Innovation Challenge - Data Science 🔗 <i>GE Digital</i>

Skills

Machine Learning <i>Supervised, Unsupervised</i>	● ● ● ● ●	Natural Language Processing <i>Basics of NLP to LLM</i>	● ● ● ● ●
Computer Vision <i>Image Classification</i>	● ● ● ● ●	Distributed Computing <i>PySpark, MapReduce</i>	● ● ● ● ●
MLOps <i>Model Deployment</i>	● ● ● ● ●	Audio Signal Processing <i>Audio Segmentation & Classification</i>	● ● ● ● ●

Technology Stack

- **Languages** : Python | PySpark | SQL | HQL
- **IDE** : JupyterLab | VS Code | PyCharm | Databricks
- **Cloud/API** : GCP | Azure | RestAPI | FastAPI
- **AI Platforms** : Google's Vertex AI | Azure AI Platform | H2O AI
- **Frameworks** : Scikit-Learn | TensorFlow | Keras | PyTorch | FastAI | LangChain | ONNX
- **Databases** : NoSQL | Relational | VectorDB | GraphDB | EmbeddingDB | Delta Lake
- **Visualization** : PowerBI | Tableau | Streamlit | Panel | Dash | Flask | Matplotlib | Seaborn | Plotly
- **Algorithms** : Bagging (Random Forest) | Boosting (XgBoost, LightGBM, Catboost) | Clustering (K-means, DBSCAN, Hierarchical) | Linear Regression (Logistic, Ridge, Lasso) | Dimensionality reduction (PCA, SVD, LDA) | AutoML | Computer Vision (ResNet, MobileNet, YOLO etc) | Natural Language Processing (TF-IDF, ELMo, BERT, RoBERTa, GPT etc) | Time Series (Sktime, Tsfresh, tslearn, Dart, Prophet etc)
- **MLOps** : MLFlow | WandB | DVC | DagsHub | Feature Store (Feast, Databricks) | TFX Serving | Docker

Professional Experience

Sep 2021 – Present
(Telecom, Network)

Lead Data Scientist [2+ Years]

Reliance Jio, Bangalore

| XGBoost | LightGBM | SHAP | NetworkX | PySpark | Time Series | LLM |

- Built a multi model architecture to predict customer churn, followed by a deep reason code analysis to retain the customer and improving telecom services.
- Built a recharge delay prediction model in order to avoid potential revenue leakage and thus in a way reduce target audience for sending sms reminders.
- Used clustering to identify suitable geolocations for sales team to setup camp for upselling sim cards & 5G Dongles.
- Fine-tuned RoBERTa model to categories users textual feedbacks received through NPS surveys into different categories.
- Built a time series predictive maintenance model to monitor and prevent telecom tower failure in advance.

Jan 2018 – Aug 2021
(Insurance, Investment)

Senior Data Scientist [3.8 Years]

M&G Prudential, Mumbai

| BERT | ResNet | LSTM | ESG |

- Developed a regression model which estimates the carbon emission values of companies who haven't disclosed it in their Annual Report.
- Worked on huge volumes of text data to identify companies intent to become carbon neutral using Google's BERT Model.
- Developed first in the segment Facial Analytics tool to access Age, Gender and BMI of an onboarding customer based on a selfie.
- Developed MISO Customer Voice matching model to track customer journey also built a Bi-Directional LSTM model to separate IVR from the speech with 90% accuracy.
- Built several models to access customer's location, occupation and lifestyle risk in a way helping customer onboarding process by dynamically asking questions based upon individual's risk profile.

Feb 2016 – Dec 2017
(Energy, Oil & Gas)

Data Scientist [2 Years]


Cognizant Technology Solutions, Chennai

| Naive Bayes | Topic Modeling | SVM |

- Built ESG Intensity Prediction Tool using unsupervised news data from lexisnexis and Topic Modeling.
- Built inhouse tool for predicting defect severity using Naive Bayes Algorithm.
- Worked on data migration project, to move data from teradata to hadoop platform for British Gas.

Jun 2014 – Aug 2014
(Medical Imaging)

Imaging Specialist - Intern [3 Months]

MIT Media Labs - Camera Culture Group, USA 




| OpenCv | MATLAB | Video Analysis |

- Selected in top 100 engineers all over India by Massachusetts Institute of Technology(MIT) for 3 month internship program named ReDx (Redesigning Diagnostics) which is related to making portable and new ways of health diagnostics for common man, conducted by MIT & Harvard in collaboration with IITb & Tata technology and design center.
- Worked on projects related to affordable health diagnosis like Measuring Blood Pressure using a smartphone (SnapBP) and High Field of view endoscope.

Education

2013 – 2015 Mtech, I&C	Master of Technology in Instrumentation & Control <i>Vishwakarma Institue of Technology, Pune</i> GPA : 8.8/10
2009 – 2012 Btech, E&TC	Bachelor of technology in Electronics & Telecommunication <i>Sinhgad College of Engineering, Vadgaon , Pune</i> GPA : 7.7/10
2006 – 2009 Diploma, E&TC	Diploma in Electronics & Telecommunication <i>Government Polytechnic, Amravati</i> GPA : 9.2/10

Certificates

Google Cloud Big Data and Machine Learning Fundamentals  <i>Coursera - Dec-2022</i>	Deep Learning in Python by Dan Becker  <i>Datacamp - Feb-2018</i>	Deploying Scalable Machine Learning for Data Science  <i>Linkedin - Jun-2021</i>
--	--	---

Interests

Travelling | Technical Blog Reading | Solving Puzzles | Watching Podcasts | Playing Cricket

Declaration

I hereby declare that all the information mentioned above in the resume is correct to the best of my knowledge and belief.