

Arkaprabha Bhattacharyya

Senior Associate, Cognizant Technology Solutions

Data Scientist with 6 years of extensive Information Technology (IT) industry experience- spanning domain areas of energy, finance, retail and telecom- along with a strong educational background. Experienced at creating data regression, classification, clustering and forecasting models, to deliver insights and implement action-oriented solutions to complex business problems. Highly skilled in Machine Learning, Natural Language Processing (NLP), Neural Networks, Statistics and Data Storytelling.



✉ arkaprabha.b92@gmail.com

📍 Bengaluru, India

📧 arkaprabha.bhattacharyya1

📞 +91 9581370658

🌐 linkedin.com/in/arkaprabha92

🐙 github.com/Arkaprabha-B

TECHNICAL SKILLS

Machine Learning Algorithms

Supervised Learning, Unsupervised Learning and Ensemble Learning

Data Visualisation

Tableau and Microsoft Excel

Text Analytics

Natural Language Processing (NLP) using spaCy and NLTK

Languages

Python and R

Neural Network (NN) and Deep Learning

Multi Layer Perceptron, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN)

Hadoop Ecosystem

Apache Hive, Oozie, PySpark (Apache Spark)

WORK EXPERIENCE

Senior Associate Cognizant Technology Solutions

04/2020 - Present

Cognizant Digital Business- Artificial Intelligence and Analytics

Bengaluru, Karnataka, India

- Developed Music Recommendation System for world's second largest smartphone manufacturer, using music metadata and consumer behaviour.
- Recommendation scenarios include- new songs from favourite artist, songs you may like, top playlist of the genre and playlists consumer may like.
- Worked as a Data Scientist to implement Time Series and multiple Machine Learning (ML) models for revenue forecasting.
- Trained and guided mentees on machine learning algorithms and techniques.
- Extensively used Python libraries Numpy, Pandas, Scipy, Scikit-learn, Tensorflow and Keras.
- Technologies: Collaborative Filtering, K- Nearest Neighbor (KNN), Time Series, Python, Hive (Hadoop Ecosystem), SQL
- Tools: Anaconda Navigator (Jupyter Notebook), PyCharm, Microsoft Excel, Linux Platform (Hadoop HDFS), Tableau.
- Team size: 15.



WORK EXPERIENCE

Technology Analyst Infosys Limited

04/2018 – 03/2020

Domain: Finance (Data and Analytics)

Bengaluru, Karnataka, India

- Worked as a Data-Scientist to understand the problem statement and provide the best possible model; post exploratory data analysis, data pre-processing and hyper-parameter tuning and optimization.
- Developed an NLP pipeline that ingests 10-K reports of various publicly traded companies and built a machine learning model to predict the long term stock performance of a company from the financial reports.
- Web-scraped 10-K financial reports filed to the United States Securities Exchange Commission (US-SEC), in order to determine financial sentiments of the companies listed on NASDAQ, using Beautiful Soup.
- Data storytelling using Tableau generated dashboard/stories.
- Deployment of the models at production level using Docker container.
- Accelerated the entire Machine Learning pipeline with effective and scalable coding on agile methodology.
- Technologies: Python 3 (including Pandas, NumPy, Plotly, Beautiful Soup), ML algorithms- Linear regression, Logistic regression, Naive Bayes, Decision Trees, K-Nearest Neighbours, K-Means, Boosting (XGBoost), Bagging (Random Forest), Text Analytics, Time Series Analysis.
- Tools: R Studio, Anaconda Navigator, Jupyter Hub, Microsoft Excel, Eclipse Oxygen IDE, Tableau, Docker.
- Team size: 4.

Employee Attrition Hackathon Infosys Limited

08/2019 – Present

Bengaluru, Karnataka, India

- Developed a Machine Learning pipeline to predict employee attrition using Random Forest, which helped the organization in planning human intervention to alleviate issues faced by the workers.
- Submission among top 5 entries for the competition (F1 Score of 93% on test data).
- Technologies: Python 3 (including Pandas, NumPy), ML algorithms- Logistic regression, Decision Trees, Random Forest.
- Tools: Anaconda Navigator, Jupyter Notebook, Microsoft Power Point.

Senior Systems Engineer Infosys Limited

01/2016 – 03/2018

Domain: Energy (Data and Analytics)

Bengaluru, Karnataka, India

- Worked as a Big-Data engineer for a Data Lake program, for the largest power utility client in the United States.
- Developed high level designs, meeting the prescribed functional and non-functional requirements; ensuring the design was as per company standards and best practices; verified and validated the design versus requirement in order to ensure that all the requirements of the client were met.
- Data profiling of large data-sets; using Oracle EDQ (Enterprise Data Quality) .
- Extract, Transform and Load (ETL) using Hive, Spark 2.x and ODI 12c.
- Developed a mail trigger mechanism on the success and failure of each module, using ODI 12c.
- Automated the migration of Sqoop and Spark batch jobs using Oozie, that reduced the effort to run the scripts manually.
- Improved responsiveness and overall performance, writing highly optimised SQL and Python codes
- Technologies: Apache Hive, Apache Spark 2.x, Hadoop, MySQL, Oozie, Sqoop, Python 3.x .
- Tools: Apache Hue, Anaconda Navigator , Eclipse Oxygen IDE, iEDPS, Jupyter Notebook, Microsoft Excel, Oracle Enterprise Data Quality (EDQ), ODI 12c, Linux Platform (Hadoop HDFS).
- Team size: 20.

WORK EXPERIENCE

Systems Engineer Infosys Limited

08/2014 - 12/2015

Domain: Retail (Loyalty Points Management)

Chennai, Tamil Nadu, India

- Worked as a back-end developer for creating batch jobs and REST APIs for a US based fashion retailer and a Middle-East based loyalty reward program.
- Created batch jobs and REST APIs for the new loyalty program using Java EE, Spring framework, Hibernate and MySQL.
- Client appreciation for data migration of existing customers into the newly created loyalty platform.
- Improved code coverage using SONAR from 30% to 80%.
- Gained experience in JIRA for defect fixing, tracking and management.
- Sole responsibility to deploy code in SIT, UAT and production environment from offshore location.
- Technologies: Java, Spring Framework, Hibernate, MySQL, REST web-services, JUnit, JIRA, Sonar, Jenkins.
- Team Size: 6.

EDUCATION

Post Graduate Program (PGP) in Data Science International School of Engineering (INSOFE)

05/2019 - 10/2019

Bengaluru, Karnataka

Bachelor of Technology (B.Tech) in Electronics and Instrumentation Engineering Gandhi Institute of Technology and Management (GITAM) University

07/2010 - 04/2014

Visakhapatnam, Andhra Pradesh

- CGPA: 9.29 out of 10
- Secured 2nd rank in the Electronics and Instrumentation Engineering batch of 2010-14.

XII, All India Senior School Certificate Examination (AISSCE), Science Stream Kendriya Vidyalaya, Central Board of Secondary Education (CBSE) Board

04/2009 - 03/2010

Farakka, West Bengal

- 83.6%.
- Physics, Chemistry, Mathematics, English and Computer Science.

X, All India Secondary School Examination (AISSE) Kendriya Vidyalaya, Central Board of Secondary Education (CBSE) Board

04/2007 - 03/2008

Farakka, West Bengal

- 91.4%.

LANGUAGES

English



Hindi



INTERESTS

Artificial Intelligence

Deep Learning

Machine Learning