# Ananthakrishnan Harikumar

**16 Copenger St, Boston MA - 02120**
**+1 857-693-8782 | harikumar.a@northeastern.edu | LinkedIn | GitHub**

## SKILLS & INTERESTS

**Languages:** Python, Java, C++, Matlab, JavaScript, SQL((PostgreSQL, MySQL)
**Web Technologies:** Spring Boot, Flask, Django, FastAPI, NodeJS, HTML, CSS, Streamlit, Jinja
**Big Data Technologies:** Apache Spark, PySpark, SparkSQL, Power BI, Tableau, Airflow, AWS, Google Cloud, Azure, Docker, Kubernetes, GitHub, Terraform, Snowflake
**ML Algorithm, Statistics and tools:** Linear Regression, Logistic Regression, Gaussian Naive Bayes, KNN, SVM, Decision Tree, Random Forest, K-means, Hierarchical Clustering, PCA, Deep Learning(RNN, CNN, Transformers), NLP, NumPy, Pandas, Matplotlib, Seaborn, Plotly, Scikit-learn, SciPy, Folium, Tensorflow, Keras

## PROFESSIONAL EXPERIENCE

**Tata Consultancy Services - Data Scientist;** Kochi(IN) *Aug 2019 – June 2022*

- Handled large amount of raw data and performed data pre-processing, feature engineering, data mining with PySpark
- Implemented different forecasting algorithms for time-series analysis and prediction of different data points related to customer experience and behavior for future periods
- Developed k-means clustering model in order to understand customer backgrounds and segment of customers based on their experience and behavior to improve existing profitable relationship and to avoid customer churn using Pyspark
- Predicted customer experience across channels by collecting and analyzing both past as well as real time customer data and providing recommendations to organizations personalized to individual customers, and improved customer retention by up to 10%

## EDUCATION

**Northeastern University | Master of Science Information Systems** | Boston, MA Sep 2022 - Apr 2024
Relevant Coursework: Data Science, Advanced Data Science, Big Data Analytics, Machine Learning in Finance, Deep Learning, Cloud Computing, Data Structures and Algorithms

**Amrita University | Bachelor Of Technology Computer Science** | TamilNadu, India June 2015 - May 2019
Relevant Coursework: Pattern Recognition, Computational Intelligence, Machine Learning and Data Mining, Natural Language Processing, Data Science, Data Structures and Algorithm, Design and Analysis of Algorithms

## PUBLICATION

“Hypergraph based clustering for document similarity using FP growth algorithm” 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 332-336, doi: 10.1109/ICCS45141.2019.9065630.

## PROJECT EXPERIENCE

**Smart ETL tool for research documents - Northeastern University** *Apr 2023 – May 2023*

- Developed a full-stack web application with FastAPI and Streamlit to extract PDFs using Springer Nature’s API, transfer them to AWS S3, implement a sentence-transformer-based recommendation engine with Pinecone vector database, and customize ChatGPT 3.5 with LangChain for enhanced book queries, reducing user search time by 78%, while administering GitHub CI/CD for automated testing and streamlined development

**Assignment Management System - Northeastern University** *Sep 2023 – Dec 2023*

- Implemented a RESTful API in NodeJS with PostgreSQL ORM for an Assignment Management System, deployed on AWS using CI/CD for seamless integration, reducing deployment time by 40% through automated resource provisioning and cost-efficient, scalable AWS deployment with Pulumi, utilizing Mocha for Continuous Integration Testing with Git Actions and YAML configuration, and employing Packer to create custom Debian Linux AMIs for optimized AWS EC2 instance deployment and efficient application delivery.

**Movie Scene Genre Prediction - Northeastern University** *Sep 2022 – Dec 2022*

- Executed web scraping of Action, Animation, and Romance movie scripts from IMSDb using BeautifulSoup and httplib2, pre-processed them with Beautiful Soup, SoupStrainer, and regular expressions, and predicted movie genres using RandomForestClassifier, Multinomial Naive Bayes, and Support Vector Machine, and compared the accuracy of each model.