

Kaushik Patil

(857) 544-8130 • patil.kau@northeastern.edu • [GitHub](#) • [LinkedIn](#) • Austin, TX

EDUCATION

Northeastern University, Boston, MA
Master of Science in Information Systems

Sep 04, 2019 - May 07, 2021
GPA: 3.51 / 4.0

University of Mumbai, Mumbai, India
Bachelor of Engineering in Electronics and Telecommunication Engineering

Sep 09, 2014 - Jun 30, 2018

EXPERIENCE

Informatica, Austin, TX, United States

Senior Technical Support Engineer, Cloud Data Governance and Catalog & Enterprise Data Catalog
Technical Support Engineer, Enterprise Data Catalog (EDC)

Oct 15, 2023 - Present
Aug 09, 2021 – Oct 15, 2023

- Resolved technical inquiries related to Informatica Enterprise Data Catalog (EDC) & Cloud Data Governance and Catalog (CDGC) Products to identify the bugs which has saved several fold efforts for R&D
- Analyzed exceptions and logs to isolate root cause analysis of **Nomad Cluster and MongoDB** implementation of the product by understanding the nuances and help in product's usability and improve end-user experience
- Assisted customers to scan resources such as **Database Management, Business Intelligence** and **Cloud** to extract the metadata, identify the profiling data (uniqueness of data), the Data Domains through Data Domain Discovery, determine value frequency, identify sensitive data and build lineage to understand the flow of data in the system.
- Created and queried tables in **MS SQL Server** and **Oracle** database using DDLs to isolate, identify and resolve errors.
- Utilized **Rest API** to fetch/modify resources and resource configuration.

RAWsb Technologies Inc, Boston, MA, United States

Jun 08, 2021 – Aug 08, 2021

Software Development Engineer Intern

- Designed Covid 19 Information Intelligent Search Assistant that can explore various use cases for people to get the most helpful and correct information directly, without finding it from overwhelming amount of data online
- Cooperated with peer development team on analysis, design, and implementation of Search Robot kernel using **C** programming language for cache data for later integration with **SparkSQL** and optimize **ML** supervised learning trained AI

Wysetek Systems Technologists, Mumbai, India

Feb 04, 2019 – Jul 31, 2019

Customer Support Representative (CSR), IBM Power Series Team

- Provided service support to RISC based **UNIX** servers in IBM Power Series team of 5 for clients in CtrlS Tier-4 datacenter
- Regulated IBM Storage Manager 10 client through **LINUX** administration with error logs after scheduling and performing backup and restore operations, server monitoring and health checks maintaining unprecedented CtrlS uptime of 99.995%

ACADEMIC PROJECTS

Medical Plan RESTful Application

Jan 2021 – Apr 2021

- Concocted **REST API** using **Spring Boot** with POST, PUT, PATCH, GET and DELETE Http methods validations and status codes
- Enqueued object in RabbitMQ queue to index object and dequeued from **RabbitMQ** queue and index data in **Elastic Server**
- Implemented search queries using **Kibana** Console to visualize and retrieve indexed data

IMDb Data Analysis and Visualization

Sep 2020 – Dec 2020

- Deployed **ETL** pipelines using **Talend** and **Alteryx** from multiple source data including relational databases and flat files
- Loaded data into BI schema to better visualization with interactive BI dashboards in **PowerBI** and **Tableau**
- Improved master job by parallelization and serialization to optimize and decrease execution time by 42%

Amazon Sports Reviews Big Data Analysis

May 2020 – Aug 2020

- Initiated **MapReduce** Algorithms through **JAVA** app for 1-click analysis for specific deliverables in minimal time
- Prepared **SQL** type scripts for rapid MapReduce operations through **Hive** and **Pig** compared to **Hadoop** by 42%
- Produced graphs and dashboards using **Tableau** for go-through analysis of deliverables for easy understanding

Sentiment Analysis of Amazon Product Reviews

Sep 2019 – Dec 2019

- Explored and classified customer sentiments into positive or negative using Natural Language Processing (**NLP**)
- Quantified different accuracies of machine learning algorithms like **Logistic Regression** 72%, **K-NN** 62%, **Naive Bayes** 87%, **Random Forest** 76%, **Decision Tree** 62% and **SVM** 80% and found Naïve Bayes most optimal algorithm method

TECHNICAL SKILLS

Languages:	SQL, Python, C++, Java
Big Data Tools:	Hadoop, Apache Hive, Apache Pig, Apache Mahout, Apache Kafka, Elasticsearch, Redis, Kibana
Databases & Servers:	MongoDB, HBase, Apache Cassandra, MySQL, SSMS, Oracle 18c, PostgreSQL, Redis
Business Intelligence tools:	ER Studio, Alteryx, Talend, PowerBI, Tableau