

# Abhin Gude

abhingude9@gmail.com || (980)-267-7254 || [Abhin Gude](#)

## EDUCATION

**New York University**, New York, NY, US

May 2023

Master of Science, Computer Science, GPA: 3.93/4

**SRM University AP**, Amaravati, AP, India

May 2021

Bachelor of Technology, Computer Science, GPA: 3.82/4

**University of Wisconsin-Madison**, Madison, WI, US

May 2020

Semester Exchange Student, Visiting International Student Program, Computer Science, GPA: 3.25/4

## EXPERIENCE

**Intelix Systems**, Data Engineer

Jan 2023 – Present

Client: **Universal Service Administrative Company**, Washington, DC, US

- Spearheaded end-to-end development and optimization of ETL workflows using Pentaho Data Integration (Kettle), orchestrating seamless data integration from diverse sources including relational databases (MySQL, PostgreSQL), cloud platforms (AWS S3, Azure Blob Storage), and real-time streaming sources (Apache Kafka).
- Employed advanced data transformation techniques within Pentaho, leveraging its extensive array of built-in functions and plugins, alongside SQL and scripting languages (JavaScript, Python), to engineer intricate data pipelines.
- Implemented sophisticated cleansing, normalization, and enrichment processes, ensuring data quality and consistency across the organization.
- Architected scalable and fault-tolerant data processing solutions, harnessing Pentaho's distributed computing capabilities and parallel processing techniques.
- Engineered optimized workflows utilizing clustering technologies (Apache Hadoop, Apache Spark) and containerization platforms (Docker, Kubernetes), enabling efficient resource utilization and accelerated data throughput.

**Computer Science and Math Department New York University**, Course Assistant, New York, NY, US

Dec 2021 – May 2023

- Course Assistant for Database Systems, Data Management and Analysis, as well as Discrete Mathematics courses, under Professors Zeidenberg, Cereste, and Esposito.
- Assisted professors in managing course logistics, including organizing materials, facilitating communication with students, conducting doubt clearing sessions and supporting instructional activities.

**MTA New York City Transit**, Data Research Associate Intern, New York, NY, US

June 2022 - Aug 2022

- Orchestrated the seamless archival of diverse data sources including Postgres, MySQL, and Oracle databases to Amazon Web Services (AWS) S3 storage, handling over 1 TB of data with 99.9% accuracy and accessibility.
- Pioneered the development of a cutting-edge production data pipeline, efficiently processing data from the mosquito message queue to SQLite and MongoDB, resulting in a 40% improvement in data transfer speed and a 100% reduction in data loss during the 5 minute time frame in midnight when job is scheduled to run.
- Devised bespoke scripts for daily subway and bus performance evaluation, encompassing wait times, service delivery, and key customer-centric metrics like Average Time To Travel (ATT), Average Journey Time (AJT), and Customer Journey Time (CJT). These efforts resulted in a notable 30% enhancement in operational efficiency, service reliability and used for open data.

**Bajaj Finserv Health Limited**, Data Scientist Intern, Pune, MH, India

Jan 2021 - Jul 2021

- Conducted exhaustive customer data analysis to fuel strategic business development and targeted marketing initiatives, particularly focusing on health insurance sales trends and precise customer segmentation for personalized outreach campaigns.
- Revamped and optimized Search API functionality, implementing automated zero search results diagnosis using Python script, Microsoft LUIS, and language processing API. Successfully slashed zero search results from 18.9% to an impressive 2.13% during the migration from Azure to Elastic search, ensuring enhanced user experience and search relevance.
- Leveraged the power of Power BI and Google Analytics Dashboards for comprehensive insights into search personalization, health insurance sales dynamics, and doctor-patient booking trends, empowering data-driven decision-making and fostering agile business growth strategies.
- Propelled user engagement and platform usage by introducing a cutting-edge medical articles search feature, resulting in a remarkable 29% surge in search bar hits and a notable 13.3% increase in customer acquisition, exemplifying a commitment to innovation and user satisfaction.

**Computer Science Department SRM University AP**, Research Assistant, Amaravati, AP, India | [code repo](#)

Jul 2020 - May 2021

- Designed and generated treatment plans for diseases like COVID and Cancer using graph techniques, Artificial Intelligence (Graph - Node Insertion, Search and Path) and Big Data (Large data stored in Hadoop File System).
- Tasks include Data Collection, Data Cleaning, Graph Modelling, Data Insertion into graphs, generating treatment plan and validating results.

## PROJECTS

### Movie Recommendation System [BigData, Spark, SQL, Python, Machine Learning]

Spring 2022

Capstone Project | Big Data Course | New York University | [code repo](#)

- Developed a movie recommendation system using big data technologies such as Apache Spark and machine learning algorithms in Python.
- Implemented collaborative filtering, content-based recommendation, and hybrid recommendation techniques to analyse user behaviour and movie attributes, resulting in personalised recommendations.
- Improved user engagement and satisfaction metrics by 81.7% through the implementation of advanced machine learning models such as Matrix Factorisation and Neural Collaborative Filtering and incorporated sentiment analysis for user reviews, leading to increased click-through rates and retention.

### Facial Snapchat Filters [Python, Deep Learning, CNN]

Spring 2022

Capstone Project | Computer Vision Course | New York University | [code repo](#)

- Engineered a Convolutional Neural Network (CNN) model tailored to train and accurately predict facial landmarks, enhancing facial recognition capabilities.
- Employed advanced affine transformations to seamlessly morph filters onto image features, refining the precision and adaptability of the model.
- Augmented input data, leveraging affine and intensity transformations to enrich the dataset of human faces. Utilized sophisticated image mapping techniques, incorporating back mapping to resolve distortions effectively.
- Integrated alpha blending techniques during the fusion of facial images and filters, ensuring seamless integration and natural appearance. Additionally, applied filters to personal facial images, demonstrating real-world applicability and versatility of the model.

### American Hospital Information and Management System [Streamlit, Python, PostgreSQL, MySQL]

Fall 2021

Capstone Project | Database Systems Course | New York University | [code repo](#)

- Designed and deployed a comprehensive hospital management system web application utilizing Streamlit and PostgreSQL Database, ensuring efficient data management and user interaction.
- Implemented essential features such as location-based hospital and doctor search functionalities, seamless access to patient's past appointments and medication history, and real-time identification of nearest hospitals equipped with ambulance facilities, enhancing user experience and accessibility.
- Successfully deployed the application on the NYU Jedi server and Streamlit platform, ensuring seamless access and optimal performance for users.

### Lake Mendota and Lake Monona Ice Cover Prediction [Python, Machine Learning, Regression, WebScraping]

Spring 2020

Individual Project | University of Wisconsin-Madison | [code repo](#)

- Developed a regression model to forecast ice cover on Madison lakes, Mendota, and Monona, providing insights into future trends.
- Conducted in-depth data analysis to uncover intriguing patterns and trends in local climate variations over time.
- Utilized Beautiful Soup for efficient web scraping of data from the Ed Hopkins Wisconsin State Climatology office website, ensuring accurate and reliable data collection for analysis.

## TECHNICAL SKILLS

<b>Programming Languages:</b>	Python, Java, C, C++, Node.js, MATLAB, PHP, R, Haskell, Julia, Scala, Hive, Spark, Hadoop
<b>Database Languages:</b>	SQL, MySQL, PostgreSQL, Oracle, MongoDB, SQLite
<b>Data Warehouses:</b>	Azure Synapse Analytics, Snowflake, Amazon Redshift
<b>ETL Tools:</b>	Pentaho, Informatica, SSIS, IBM Datastage
<b>Analytics:</b>	Power BI, Tableau, Google Analytics, Databricks

## HONORS AND AWARDS

- |   |          |
|---|----------|
| • Master's Degree Scholarship of \$16000 from New York University   | Sep 2021 |
| • Winner of Optum Stratethon Season 2 - Living our values award   | Dec 2020 |
| • Awarded bounty prize by Maker DAO of 12 ETH for the Best Governance Hack at ETH Denver, CO, USA   | Feb 2020 |
| • Director's Scholarship of \$14500 from SRM University AP and \$5000 from University of Wisconsin for semester exchange program at University of Wisconsin-Madison | Dec 2019 |
| • Winner of TechQuiz during TechFest at SRM University AP   | Apr 2018 |
| • Bachelor's Degree Scholarship of \$13500 from SRM University AP for entire CS Bachelor's Degree   | Aug 2017 |

## ASSOCIATIONS AND MEMBERSHIPS

- |   |          |
|---|----------|
| • Member of NYU Machine Learning and Business Analytics Club at New York University | Jan 2022 |
| • Member of ICF, Bridges and Spades at New York University                          | Sep 2021 |
| • Member of Coding for Good Club at University of Wisconsin-Madison                 | Jan 2020 |
| • Member of the Student Council at SRM University AP                                | Jul 2018 |