



Sumit Mishra

[LinkedIn](#) | [Github](#)

GET IN CONTACT

Mobile: +91-7400255701 / 8689907119

Email: sm27598@gmail.com

PERSONAL DETAILS

- Total Experience 5 Years 2 Months
- Current Location Mumbai
- Date of Birth May 27, 1998
- Gender Male
- Marital Status Single / Unmarried

TECHNICAL SKILLS

- Python
- Natural Language Processing
- Machine Learning
- Large Language Model (LLM)
- Deep Learning
- Neural Networks
- Pytorch
- Flask
- Computer Vision
- Apache Airflow
- SQL
- Apache Kafka
- Pyspark
- Keras
- Data Processing
- Feature Engineering
- Data Visualization
- Exploratory Data Analysis
- Scikit-Learn
- Pandas
- Numpy
- Spacy
- Matplotlib
- Prompt Engineering
- Web Technologies
- HTML
- CSS
- Javascript
- JQuery
- PHP
- ASP.NET

WORK EXPERIENCE

Jan 2022 to Present

Research Assistant (Machine Learning & Natural Language Processing) at University of Utah

- Spearheaded advanced research and development (R&D) projects spanning natural language processing (NLP), data engineering, and machine learning (ML), contributing to the optimization of algorithms and models.
- Led data collection efforts, gathering extensive unstructured and structured data from diverse web sources to construct high-quality datasets crucial for model training and fine-tuning.
- Designed, implemented, and fine-tuned deep learning models in PyTorch to address various applications, including computer vision, NLP, and data analysis, ensuring robust and efficient solutions.
- Leveraged cutting-edge language models such as BERT, BART, and GPT in Huggingface and in native PyTorch for a range of NLP projects, pushing the boundaries of language understanding and generation.
- Developed end-to-end ML models and pipelines, covering data acquisition, preprocessing, feature engineering, model training, and deployment.
- Actively engaged in knowledge-sharing sessions, workshops, and conferences to stay at the forefront of AI developments, promoting continuous learning and innovation.
- Collaborated closely with peers and mentors to refine research methodologies, experimental designs, and project execution strategies, fostering a collaborative and productive research environment.
- Conducted comprehensive reviews of NLP-related research papers, translating their insights into practical applications within ongoing projects, driving innovation.
- Demonstrated exceptional analytical skills, adeptly interpreting research results and extracting actionable insights to inform future investigations and project directions.
- Maintained meticulous project documentation to ensure transparency, replicability, and efficient knowledge transfer among team members.

May 2021 to May 2022

Data Engineer at MentorStudents.org

- Designed and implemented automated data pipelines using Apache Kafka in Docker and Apache Airflow for efficient data handling and processing, streamlining data

- C#

SOFT SKILLS

- Communication
- Research and Development
- Problem Solving

LANGUAGES KNOWN

- English
- Hindi
- Marathi

COURSES & CERTIFICATIONS

- Applied Machine Learning
- Python For Machine Learning, Statistics For Data Science & Machine Learning, Data Visualization, Etc.

AWARDS

- **1st Rank** in 3P (Project-Prototype-Presentation) Evaluation Event, awarded by B.Sc. (I.T.) Association.
- **3rd Rank** (81.6%) in Semester 5 University Examination for the Year 2017-18, awarded by V.E.S. College.
- **1st Rank** (145/150) in the Advanced Java University Examination for the Year 2017-18, awarded by V.E.S.College.
- **1st Rank** in Quiriosity Event of Vihaan'18, an Inter-Collegiate Technical Fest.
- **2nd Rank** in Iotics (Internet of Things) Event of Vihaan'18, an Inter-Collegiate Technical Fest.

VOLUNTEER & EXTRA-CURRICULAR ACTIVITIES

- Organized an event on cyber-security for college students.
- Volunteered for organizing a technical fest of the B.Sc. (I.T.) department of V.E.S. college.
- Helped and Participated in Annual Extension Activities (Lifelong learning & Extension) related to Information Technology, Industry Orientation, Entrepreneurship & Career Guidance, etc.
- Engaged in various inter-collegiate technical competitions and disaster awareness programs, showcasing a commitment to learning and community service.

workflows.

- Achieved scalability and fault tolerance for handling large data volumes, ensuring consistent data flow and reliability in data processing.
- Demonstrated expertise in natural language processing (NLP) by effectively segmenting text into sentences and applying multi-label and multi-class classification techniques, contributing to advanced data analysis.
- Developed and integrated machine learning models into data pipelines, enabling automated data-driven decision-making and enhancing predictive capabilities.
- Created interactive data visualization dashboards to facilitate real-time monitoring and analysis, providing valuable tools for data interpretation and effective communication of insights.
- Implemented and optimized ML pipelines for model training, evaluation, and deployment, enhancing the organization's data-driven capabilities.
- Developed real-time and batch data streaming systems for dynamic data sources, enabling continuous monitoring of trends and providing valuable insights for analysis and research.
- Enhanced data-driven decision-making by automating daily processing of external data sources, ensuring timely insights, and supporting informed decisions.
- Collaborated with cross-functional teams to optimize data engineering processes, fostering knowledge sharing and continuous improvement of data accuracy.

Mar 2021 to May 2021

Data Science and Business Analytics Intern at The Sparks Foundation

- Wrote code to collect, process, and analyze data from both internal and external sources, enhancing data-driven decision-making.
- Developed and fine-tuned machine learning models using Python and scikit-learn, improving predictive capabilities.
- Successfully completed various projects in Machine Learning, NLP, Deep Learning, and Computer Vision.
- Leveraged skills in Keras for deep learning projects, enhancing model accuracy and performance.
- Gained hands-on experience in data analysis and natural language processing (NLP), contributing to data-driven insights.

Sep 2019 to Mar 2021

Software Developer at TalkValley LLC

- Developed responsive web applications, ensuring user-friendly interfaces and optimal functionality.
- Collaborated cross-functionally to gather requirements and create effective web solutions.
- Integrated third-party APIs to enhance application functionality, improving user experience.
- Conducted rigorous testing and debugging to maintain seamless application performance.
- Successfully maintained and updated existing web applications, staying current with emerging technologies.

- Ensured cross-browser compatibility, delivering consistent user experiences across platforms.
- Contributed to code quality through active participation in code reviews and provided technical support for issue resolution.

Aug 2018 to Sep 2019

Programmer at Tata Consultancy Services

- Applied operating system and program product fixes regularly, ensuring the stability and reliability of systems.
- Ensured strict adherence to process conformity and consistently achieved SLAs (Service Level Agreements).
- Managed network devices, servers, batches, and jobs effectively during transitions to prevent irrelevant and erroneous issues.
- Collaborated with various mainframe services to address specific requests and optimize system performance.
- Liaised with software vendors to report, track, and efficiently resolve issues related to software products, ensuring seamless operations.
- Coordinated external dependencies with neighboring teams and subject matter experts (SMEs), facilitating smooth project execution.

EDUCATION

Post Graduation

Jan 2023 to Present

M.Sc. (Data Science)

Chandigarh University

Graduation

Jun 2015 to May 2018

B.Sc. (Information Technology)

University of Mumbai

“A” Grade (80%)