

Sumit Mishra (He/Him/His)

Chembur, Mumbai, Maharashtra 400074, India

sm27598@gmail.com

(+91) 7400255701, (+91) 8689907119

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

SKILLS

- **Programming and Development:** Python, Flask, ASP.NET, C#, Core Java, HTML, CSS, JavaScript, PHP.
- **Machine Learning and Artificial Intelligence:** Natural Language Processing (NLP), Deep Learning, Computer Vision, Machine Learning Algorithms, Data Visualization, Keras, LLM (Large Language Model), Apache Spark, PySpark, Apache Airflow, Apache Kafka.
- **Data Engineering & ETL:** Data Analysis, Data Preprocessing, Feature Engineering, Data Visualization, Data Cleaning, Exploratory Data Analysis, Database Management Systems, SQL.
- **Tools & Frameworks:** JQuery, TensorFlow, PyTorch, Keras, Scikit-Learn, Pandas, NumPy, Spacy, Matplotlib, Seaborn, MySQL, Filezilla, Docker.
- **Additional Skills:** Communication, Problem-Solving, Research and Development (R&D), Windows Server, Unix, Mainframe Programming, Prompt Engineering.
- **Languages:** Professional working proficiency in English, Native or bilingual proficiency in Hindi, and Limited working proficiency in Marathi.

PROFESSIONAL EXPERIENCE

University of Utah

Salt Lake City, Utah, USA (Jan 2022 - Present)

Research Assistant (Remote)

- Successfully executed multiple NLP projects, demonstrating project management skills and innovative problem-solving abilities.
- Proficient in working with various NLP models, including fine-tuning and optimizing them for specific tasks, showcasing deep model knowledge.
- Conducted research on diverse NLP techniques, emphasizing performance analysis and efficient model implementation.
- Implemented transfer learning techniques, achieving high accuracy and showcasing the potential of transfer learning in NLP applications.
- Committed to knowledge sharing by providing code resources, enhancing accessibility for fellow researchers and practitioners.
- Demonstrated adaptability by adjusting project directions and actively seeking solutions to overcome challenges, displaying a proactive problem-solving approach.
- Here are a few project details.
 - **BloggerGPT: Blog Writing AI Bot**
 - Developed an AI-powered blog writer using ChatGPT and LangChain, automating content creation.
 - Demonstrated proficiency in integrating AI technologies for content generation.
 - Achieved synergy between ChatGPT and LangChain for high-quality blog posts.
 - **BERT in Sentence Similarity Research: A Comprehensive Review**
 - Conducted an extensive examination of sentence similarity within the realm of BERT models and labeled data.
 - Explored fine-tuning as a central theme, providing valuable perspectives on optimizing these models.
 - Contributed code resources to facilitate the dissemination of knowledge in this field.
 - **Research Paper: "Efficient Attentions for Long Document Summarization"**

- Implemented the "Efficient Attentions" model, optimizing it for long document summarization.
- Proficiently handled data preprocessing, model architecture, and GPU resources.
- Achieved in-depth understanding and fine-tuning of the model's parameters.
- To learn more about the projects undertaken, please visit this link:

MentorStudents.org

Salt Lake City, Utah, USA (May 2021 - May 2022)

Data Engineer (Remote)

- Designed and implemented real-time data pipelines using Apache Kafka, Apache Airflow, and Docker for scalability and fault tolerance.
- Proficient in sentiment analysis, advanced NLP techniques, and multi-label/multi-class classification for social media data.
- Developed real-time data streaming systems for dynamic sources, creating efficient workflows and interactive data visualization dashboards for trend analysis.
- Here are a few project details.
 - ***StreamFlow: Automated Daily Data Pipeline***
 - Created an automated data pipeline, enabling daily data ingestion and processing.
 - Configured Dockerized Apache Kafka for efficient data streaming, while ensuring scalability and fault tolerance for handling large data volumes.
 - Improved data-driven decision-making by delivering timely insights from external sources through automated daily data processing.
- To learn more about the projects undertaken, please visit this link:

The Sparks Foundation

(Mar 2021 - May 2021)

Data Science & Business Analytics Intern (Remote)

- Wrote codes to collect, crunch, and analyze data from internal and external sources.
- Built and tuned various machine learning models using Python, and scikit-learn.
- Created linkages between various data to enable predictive modeling and trend analysis on a given dataset.
- Built various Machine Learning, NLP, Deep Learning, and Computer Vision-based projects.

TalkValley LLC

Mohali, Punjab, India (Sep 2019 - Mar 2021)

Web Developer (Remote)

- Developed responsive web applications using ASP.NET, HTML, CSS, JavaScript, and PHP.
- Collaborated with cross-functional teams to gather requirements and create user-friendly interfaces.
- Implemented third-party APIs to extend functionality and optimized applications for performance.
- Conducted testing and debugging, ensuring a seamless user experience.
- Maintained and updated existing web applications, continuously improving skills and staying updated on emerging technologies.

TATA Consultancy Services

Chennai, Tamilnadu, India (Aug 2018 - Sep 2019)

Programmer

- Regularly applied operating system and program product fixes.
- Ensured adherence to process conformity and achievement of SLAs.
- Effectively managed network devices, servers, batches, and jobs during transitions to mitigate irrelevant and erroneous issues.

- Collaborated with various mainframe services to address specific requests.
- Liaised with software vendors to report, track, and resolve issues related to software products.
- Coordinated external dependencies with neighboring teams and subject matter experts (SMEs).

EDUCATION

Chandigarh University

Master of Science in Data Science

Chandigarh, Punjab, India (Jan 2023 - Present)

University of Mumbai

Bachelor of Science in Information Technology

Mumbai, Maharashtra, India (June 2015 - May 2018)

- Capstone Project: "Theft Prevention Using PIR Sensor (IoT project)"
- Grade: A

CERTIFICATIONS

- Applied Machine Learning
 - By Applied AI Course (Feb 2020 - Feb 2021)
 - Details: Completed comprehensive coursework including 30+ assignments and two capstone projects focused on NLP and Computer Vision.
- Data Visualization Using Python
- Statistics for Data Science
- Data Science with Python
- Python for Machine Learning
- Statistics for Machine Learning
- Machine Learning with Python
- Machine Learning Foundations
- Introduction to Neural Network & Deep Learning
- Computer Vision Essentials
 - By Great Learning (Mar 2020 - Apr 2021)

AWARDS

- **1st** Rank in 3P (Project-Prototype-Presentation) Evaluation Event, awarded by BSc (I.T.) Association.
- **3rd** Rank (81.6%) in Sem 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 BSc (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.