Shivesh Singh

+91 7392098513 | shivesh235.github.io/bio | shivesh235@gmail.com | in/shivesh235 | github.com/shivesh235

Education

Indian Institute of Technology, Mandi
Expected May 2025

M.Tech Research in **Machine Learning**

Relevant Coursework: Pattern Recognition, Matrix Theory, Probability and Random Processes, Deep Learning, Generative AI

REC Kannauj
2019-2023

B.Tech in **Computer Science & Engineering**

Coursework: Data Structures, Algorithms, Database Management Systems, Computer Vision, Machine Learning, Deep Learning

Professional Experience

Data Science Intern
June 2022 – August 2022

Aspirevision Tech Education

- Developed mathematical models to derive predictive insights, fine-tuned ML algorithms, and analyzed model explainability for actionable outcomes.
- Conducted model evaluations with thorough analysis, ranking model responses, and providing detailed rationales for decision-making.
- Designed interactive dashboards to visualize key metrics, enhancing decision-making for business insights.
- Delivered code reviews and feedback, ensuring code quality, and actively participated in team code standard improvements.
- **Technical Stack:** Python, HTML/CSS, PyTorch, Scikit-Learn, Tableau

Projects

Sound Scene Synthesis
shivesh235/sound_scene_synthesis

- Implemented high-performing sound synthesis models, achieving top performance in the DCASE 2024 challenge.
- Provided in-depth evaluation and ranking for model outputs, focusing on quality and relevance.
- Developed 'Muzic Maesters' Web-UI, enabling interactive generation based on user inputs, showcasing strong communication and design abilities.
- **Technical Stack:** AudioLDM, Flask, VAE, U-Net, Hugging Face

Articles Research Tool
shivesh235/articals_bot

- Built a news research tool utilizing semantic search to retrieve detailed answers based on article content.
- Created an intuitive user interface, with clear explanations, enhancing accessibility for non-technical stakeholders.
- **Technical Stack:** RAG, LLM, Sentence Transformers, Langchain, Faiss

Rerender A Video
shivesh235/Rerender_v2v

- Text-guided video-to-video translation framework that adapts pre-trained image diffusion models for generating temporally consistent and visually high-quality videos.
- Hosted the app on a remote server, allows users to upload videos with a text prompt, customize advanced options like resolution and denoising strength, and apply enhancements such as super-resolution or night-light adjustments to keyframes, delivering an intuitive and feature-rich video processing experience.
- **Technical Stack:** ControlNet, GMFlow, MIRNet, NinaSR, AudioLDM, YOLO, Diffusion Models.

Technical Skills

Languages & Tools: Python, SQL, MongoDB, Scikit-Learn, PyTorch, Tensorflow

Specializations: Machine Learning, Deep Learning, Predictive Analytics, NLP, LLM, Generative AI

Publications

Under Review

- **Shivesh Singh**, Nitu Kumari and Aditya Nigam. "Deep spatio-temporal modeling of vegetation index and population density in Himachal Pradesh, India", Communicated to *Spatial Information Research*, 2024.

Achievements

- 4th Rank in DCASE Challenge 2024, showcasing expertise in model evaluation and optimization.
- AI Innovation Hackathon, Yamaha Motor Solutions, highlighting collaboration and problem-solving skills.