JIBEK GUPTA

jibek.gupta@bison.howard.edu | (202)961-3129 | linkedin.com/in/jibekgupta | github.com/jibekgupta

EDUCATION

HOWARD UNIVERSITY Washington, DC
Bachelor of Science Expected May 2026
Major in Computer Science
Relevant Coursework: Data Structures & Algorithms, Applied Data Science, Software Engineering, Database Systems, Object-Oriented Programming

TECHNICAL SKILLS

Programming Languages: Python, SQL, HTML, CSS
Machine Learning & AI: Scikit-Learn, TensorFlow, PyTorch, NLP, LLMs, RAG , Feature Engineering, Predictive Modeling
Data Science & Analytics: Pandas, NumPy, Matplotlib, Seaborn, ETL Pipelines, Statistical Analysis, Data Visualization
MLOPS & Cloud: Hugging Face, ChromaDB, AWS, Model Deployment
Tools: Git, GitHub, Jupyter Notebook, VS Code

WORK EXPERIENCE

University of Chicago, Data Science Institute Chicago, IL
Data Science Research Assistant Jun 2025 - Aug 2025
- Developed and implemented SeedBot using Retrieval-Augmented Generation (RAG) pipeline with transformers and ChromaDB, improving retrieval relevance by 35%
- Engineered and indexed 180+ multilingual legal documents, enabling broader access to legal information through advanced NLP techniques
- Designed and executed ROUGE-based evaluation benchmarks, demonstrating 25% performance gain over baseline LLMs through rigorous testing

Howard University, Historic Workshop (REU Program) Washington, DC
Data Analysis Research Assistant Aug 2024 - Present
- Analyzed and processed 100,000+ U.S. Census records (2009-2023) using Python and statistical methods to identify demographic shifts for policy recommendations
- Built and deployed interactive Shiny dashboard with dynamic year/state filters, increasing user engagement and real-time data exploration by 40%
- Presented research findings and delivered technical presentations at Howard's Research Symposium and Civic Tech DC to 200+ attendees

PROJECTS

Census Demographics Dashboard | Python, Shiny, Pandas, Data Visualization
- Developed full-stack interactive dashboard visualizing state-level demographic trends, resulting in 40% increase in stakeholder engagement
- Implemented automated data pipelines with exportable functionality, reducing manual analysis time and improving accuracy
- Designed dynamic filtering system and user interface, achieving 35% improvement in user experience scores

Machine Learning Movie Recommender System | Python, Scikit-Learn, Tkinter, NLP
- Built content-based recommendation engine using TF-IDF vectorization and cosine similarity algorithms, improving recommendation relevance by 20%
- Developed user-friendly GUI interface using Tkinter, reducing user feedback time by 35% and enhancing user experience
- Optimized feature extraction algorithms for 87,000+ movie dataset, reducing computational overhead by 15% for real-time performance

Automated Web Scraping Application | Python, BeautifulSoup, Selenium
- Engineered automated web scraping solution for 100+ product pages, reducing manual data collection time by 40%
- Developed robust scripts for dynamic content handling with structured CSV export functionality and error handling
- Optimized scraping performance and implemented scalable processing architecture with minimal latency