Vishnu Raja ❖ **Email:** vishnu1234raja@gmail.com ❖ **Phone:** +19342289706 ❖ **Location:** Stony Brook, NY

❖ **LinkedIn:** linkedin.com/in/vishnu-raja-637ba8183/ ❖ **GitHub:** github.com/VishnuRaja98

EDUCATION

Stony Brook University **Aug. 2023 - May. 2025**
Masters - MS Computer Science *Stony Brook, NY*
- Natural Language Processing, Machine Learning, Computer Vision, Reinforcement Learning, Distributed Systems.
- **Leadership:** President of Indian Graduate Student Association. **GPA 3.96**

Pune Institute of Computer Technology **July. 2016 - May. 2020**
Bachelors- BE Computer Science *Pune, India*
- Data Structures and Algorithms, OOP, Computer Graphics, AI, High Performance Computing **GPA 8.76/10**

SKILLS

AI/ML: PyTorch, TensorFlow, Hugging Face, OpenAI Whisper, Scikit-learn, NLP, LLM Research, Transformer Architectures
Programming: Python (Expert),C++, Go, JavaScript Frameworks, TypeScript, SQL, MongoDB
Tools: GCP, Docker, Kubernetes, Gitlab, Jenkins, Agile Development, JIRA, Unity, Android Studio

WORK EXPERIENCE

Vaultize Systems **Sep. 2022 - Aug. 2023**
Team Lead *Pune, India*
- Led a team of 5 to implement changes to a Digital Rights Management Platform.
- Built a Natural Language Understanding (NLU) pipeline for semantic file search using descriptive prompts.
- Designed and developed a new File Organization system with advanced features for reducing file retrieval time.

ElasticRun **Aug. 2020 – Sep. 2022**
Data Scientist, SDE 2 *Pune, India*
- Developed scalable backend systems (Python/Flask) and frontend PWAs (Angular/Svelte) serving 1M+ customers.
- Created a Random Forest Classifier that accessed credit history of customers and provided credit score.
- Built a Recommendation Engine using Correlation analysis and Clustering techniques.
- Deployed Docker and Kubernetes configurations, ensuring high availability and scalability.
- Automated reconciliation of 2,000+ transaction discrepancies and performed root cause analysis (RCA).
- Participated in the full SDLC including code review, source controls management on GitLab, Unit and A/B testing.

PROJECTS.

Audio Transcription for speech impaired **May. 2024 - Current**
- Fine-tuned OpenAI Whisper using PyTorch and Hugging Face Transformers to transcribe speech for individuals with Dysarthria, a condition common in Parkinson's patients to create general use and user specific models.
- Achieved a 50% improvement in Word Error Rate (WER) with general models and an average of 80% improvement with user-specific models. A novel two-step training method provided an additional 10% improvement in WER.
- Currently working on a live transcription method for these models using an open source Whisper-live project.

Hyperbolic Embeddings in Machine Learning **May. 2024 - Current**
- Researched and implemented hyperbolic embeddings to optimize representation of graph-structured data, achieving state-of-the-art performance with 32 dimensions (vs. 786 in Euclidean space).
- Applied hyperbolic geometric transformations to tasks like paraphrase detection and large-scale graph-based systems.
- Technologies: PyTorch, Hugging Face Transformers, Hyperbolic Embeddings, Graph Neural Networks (GNN.)

Blackjack Counting Cards using Reinforcement Learning link **Dec. 2024**
- Applied Deep Q-Learning and SARSA to train an agent in a customized OpenAI Gymnasium Environment.

Fail-Stop & Byzantine Fault-Tolerant Distributed System Transaction Processing System link **Nov. 2024**
- Built modular and maintainable code in Go for Paxos and Linearized PBFT (Practical Byzantine Fault Tolerance.)
- Enabled transaction Linearizability and Serializability by means of 2 Phase Commit and 2 Phase Locking.
- Benchmarked transaction achieving throughput of 80 transactions per second and an average 15 ms latency.