

# PRATIKSHA PRADEEP SHARMA

(352)-870-6364 | GitHub: <https://github.com/PratikshaSharma4> | LinkedIn: <https://www.linkedin.com/in/pratiksha-sharma-40488115a/>

Email: [pratikshas2343@gmail.com](mailto:pratikshas2343@gmail.com) | Gainesville, FL

## EDUCATION

**Master of Science, Agriculture, and Biological Engineering- Information Systems** | University of Florida *Aug 2021 - Dec 2023*  
*Coursework: Statistical Machine Learning, Bayesian Analysis for Machine Learning, GPA – 3.67*

**Master of Science, Electrical and Computer Engineering** | University of Florida *Aug 2019 - Dec 2021*  
*Coursework: Neural Networks and Deep Learning, Fundamentals of Machine Learning, GPA – 3.63*

**Bachelor of Science, Electronics, and Telecommunications Engineering** | University of Mumbai *Jul 2015 - May 2019*

## SKILLS

- **Languages:** Python, SQL, R, MATLAB, Java, C++
- **Machine Learning Frameworks:** Keras, TensorFlow, Caffe, Scikit-Learn, SpaCy, PyTorch, PySpark, NLTK, Spark, Docker, Kubernetes
- **Data Analysis & Visualization:** Pandas, NumPy, Excel, Matplotlib, Seaborn, ggplot2, Tableau, PowerBI
- **Cloud Platforms:** Amazon Web Services (AWS), Microsoft Azure, Snowflake
- **Statistical Methods:** Statistical Analysis, Hypothesis Testing, ANOVA, Time Series Forecasting, Principal Component Analysis (PCA)
- **Databases:** MySQL, PostgreSQL, NoSQL
- **Other:** Clustering, Sentiment Analysis, supervised, and unsupervised learning

## EXPERIENCE

**Data Science Intern** | MindHome Inc. *Aug 2023 –Present*

- Leveraged Python and SQL to extract valuable insights from vast robotic data repositories, driving informed decision-making.
- Applied machine learning, data analysis, and statistical modeling on sensor data to predict potential issues, resulting in a 15% performance enhancement.
- Detected anomalies using advanced statistical methods (hypothesis testing, ANOVA, time series), with 98% accuracy.
- Engineered end-to-end ML pipelines with Azure ML, Python, and SQL. Reduced model development time by 40% through the application of AutoML for algorithm exploration and hyperparameter tuning.

**Graduate Research Assistant** | University of Florida *Mar 2021 – May 2023*

- Utilized advanced statistical techniques and predictive modeling (**Logistic Regression, XGBoost, Random Forest, MLP, DNN**) to estimate cow pregnancy probabilities, effectively handling imbalanced data with 29% positive outcomes.
- Conducted comprehensive exploratory data analysis, preprocessing, and feature engineering on a dataset of 26,000 cow records.
- Improved model performance by 15% through effective outlier handling and feature selection methods.
- Implemented hyperparameter tuning and 10-fold cross-validation, achieving impressive results, including an AUC of 0.85 and an F1 score of 0.87 using the XGBoost classifier.

## PROJECTS

### Regeneration of Images and text using generative learning and deep neural networks

- Developed Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) in Python, using Keras, and TensorFlow, to generate MNIST digits.
- Evaluated the generated digits against manually labeled data, achieving an 86% accuracy using a CNN classification model.
- Trained an LSTM model on a dataset comprising 50,000 IMDB movie reviews to generate coherent and contextually related sentences based on user-provided seeds.

### Movie Recommender System

- Implemented a content-based recommender system using cosine similarity on the TMDB 5000 movies dataset, incorporating Python, Scikit-Learn, and Natural Language Processing (NLP) techniques, including NLTK and CountVectorizer.
- Developed an interactive web application using Streamlit, and deployed it on Heroku, enabling users to access personalized movie recommendations based on their preferences.

### Customer Analytics for Business Growth

- Employed a combination of Hierarchical and K-means clustering techniques, enhanced by Principal Component Analysis (PCA) for customer segmentation, reducing dimensionality by 30%.
- Conducted comprehensive descriptive analysis by brand and segment, uncovering valuable consumer behavior patterns.
- Utilized linear and logistic regression models to drive elasticity modeling, resulting in a substantial 20% improvement in purchase probability, 15% higher brand choice prediction accuracy, and a 25% increase in purchase quantity estimation.
- Achieved over 90% accuracy in predicting future customer behavior using an Artificial Neural Network (ANN).

### Rock Paper Scissor Game using OpenCV

- Created a real-time Rock, Paper, Scissors game using OpenCV, and Tensorflow for hand sign recognition and game output prediction using a custom dataset using data augmentation (rotation, scaling, and translation) comprising thousands of labeled hand-sign images.
- Fine-tuned the NASNETMobile architecture with a focus on optimizing for both accuracy and inference speed, leveraging transfer learning from ImageNet weights.

## CERTIFICATIONS

- Python for Data Science and Machine Learning
- The Ultimate MySQL Bootcamp