Sreeja Kodati

 Sreeja1391   in sreejakodati 📞 (945)-249-1998 ✉ skodati@wisc.edu

EDUCATION

University of Wisconsin - Madison
Master's in Statistics and Data Science   *Sep 2022 – May 2024 (CGPA: 3.6/4.0)*
- **Awards:** Graduate Scholarship for academic achievement.
- **Courses**: *Statistical Data Visualization, Statistical Methods, Statistical Consulting, Text Analytics and Business Application, Statistical Learning Theory, Deep Learning and Generative Models, High-Performance Computing.*

ICFAI Foundation for Higher Education
Bachelor's in Computer Science Engineering   *Aug 2018 – Jun 2022 (CGPA: 9.25/10)*
- **Awards:** Merit-based scholarship for demonstrating exceptional academic performance.
- **Courses**: *Data Visualization, SQL and Database Applications, Database Management Systems, Introduction to Data Science, Data Structures, Introduction to Algorithms, Theory of Computation.*

WORK EXPERIENCE

Data Analyst, Center for European Studies   *Aug 2023 – Present*
- Analyzed 10+ years of fellowship recipients data using statistical methods such as linear and logistic regression in R to provide valuable insights on various aspects of the fellowship program.
- Assisted in website maintenance and updates which involves integrating data visualization elements with HTML, CSS and WordPress customization.
- Performed accurate data compilation for research grant applications which employs several data cleaning techniques and validation methods.
- Developed interactive dashboards using Tableau for enhancing marketing strategies within the international division.

Full Stack Developer, Bidgala, Canada - *[Remote Internship]*   *Jan 2022 - Jun 2022*
- Developed and deployed a data-driven application using Pyscript and Django, enhancing customer engagement by 45% through personalized newsletters and a "recently viewed" section in the user interface.
- Implemented Lazy Loading technique in Django templates for images, resulting in a 40% faster page load time.
- Collaborated with senior software engineers to maintain, troubleshoot, and resolve application-based issues/bugs.

PROGRAMMING SKILLS

Programming and Scripting Languages: Python, R, Java, C, C++, Dart, SQL, HTML, CSS, Javascript, Bash
Frameworks: Django, Hadoop MapReduce, PySpark, Scikit-learn, TensorFlow, Keras, PyTorch, LangChain, VectorDB
Data visualization and web development: Shiny(R), D3.js, Tableau, Android Studio

ACADEMIC PROJECTS

Generalized Transfer Learning Pipeline for Pretrained Image Datasets Github
- Conducted an in-depth analysis to understand the impact of fine-tuning earlier versus final layers on transfer learning for image classification and experimented with VGG16 and ResNet18 on CIFAR10 and Fruits360.
- Resulted in an observation that freezing initial layers and finetuning the rest has improved the model's performance, demonstrating the trade-off between feature transferability and dataset-specific feature learning.
- Visualized features from different experiments which showed that earlier layers capture low-level features and final layers capture high-level details, and thus freezing the former improves the overall performance.

Analyzing Sephora Skincare Products and Reviews with NLP techniques Github
- Addressed questions pertaining to product popularity, top-reviewed products, patterns in customer feedback and preferences by using Random Forest Regressor model, NER techniques and K-means clustering algorithm.
- Developed an interactive and customized FAQ retrieval chatbot interface utilizing the ChatGPT API, LangChain and VectorDB which answers queries from users using the set of provided documents.

Extracting popularity insights for Spotify data with Statistical Analysis and Parallel Computing Github
- Implemented various machine learning models such as Linear Regression, Random Forest Regressor, and MLP Regressor to predict song popularity using Spotify's defined set of 18 audio features.
- Conducted in-depth analysis on 45 genre-specific data files utilizing HTCondor's parallel computing techniques to identify top correlated features within each genre, using each track's metadata analysis.

GreenEats: Food recommendation system with Named Entity Recognition (NER) Github
- Employed NLP techniques such as NER to extract dishes from Yelp reviews dataset and created a web application using Shiny to recommend the best vegetarian and vegan dishes in Philadelphia, enhancing the food experience for plant-based food enthusiasts.