



Thota Vivek

Roll No.:S20230010244

Computer Science Engineering

Indian Institute Of Information Technology, Sri City

+91-9676373804

thota.vivek2005@gmail.com

vivek.t23@iiits.in

Github | thota-vivek05

EDUCATION

| Degree/Certificate | Institute/Board | CGPA/Percentage | Year |
|--------------------|--|-----------------|-------------|
| B.Tech. (CSE) | Indian Institute of Information Technology, Sri City | 8.41 | [2023-2027] |
| Senior Secondary | Sri Chaitanya Junior College | 975/1000 | [2023] |
| Secondary | St. Peters High School | 85.5% | [2021] |

PROJECTS

•Telugu News Classification using N-gram Model

Tools: Python, NLTK, Scikit-learn, Pandas, Jupyter Notebook

- Built a text classification system to categorize **10,000 Telugu news articles** across five categories: business, sports, nation, entertainment, and editorial.
- Preprocessed the dataset using tokenization, stop-word removal, stemming, and normalization to handle linguistic complexities of Telugu text.
- Implemented N-gram models (Unigram, Bigram, Trigram, etc.) for feature extraction and applied Multinomial Naive Bayes for classification.
- Evaluated the model using precision, recall, F1-score, and confusion matrix, achieving **93% overall accuracy** with strong performance across all categories.
- Conducted comparative analysis with SVM and Random Forest classifiers, confirming robustness of the Naive Bayes + N-gram approach.

SKILLS

- **Programming Languages:** C, Java, Python, MySQL
- **Web Technologies:** HTML, CSS, JavaScript
- **Data Analysis:** Pandas, Numpy, Matplotlib
- **Machine Learning:** Scikit-learn
- **Tools Platforms:** Git, Jupyter Notebook, VS Code
- **Soft Skills:** Problem-solving, Team collaboration, Time management

CERTIFICATIONS

- Coursera , Artificial Intelligence
- Coursera , Complete Web Development
- Coursera , Machine Learning