

Across The Globe

Named Entity Recognition (NER) and Feature Engineering

Analyzing Popularity and Accuracy with Random Forest

Novel Kathor

Date of Submission (06/01/2025)

**Table of Contents**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| No | **Objectives** | **Page No** |
| 01 | Overview and Goals | **01** |
| 02 | Data Collection: Utilize the provided dataset of news articles. | **02-03** |
| 04 | Text Preprocessing: Conduct necessary preprocessing steps on the text. | **04-05** |
| 04 | Named Entity Recognition (NER) and Generating Synthetic Engagement Metrics | **06-08** |
| 05 | Feature Engineering: Combine entity counts with additional features to create a comprehensive feature set and Balanced the dataset. | **10-15** |
| 06 | Predictive Modeling: Training and evaluate predictive model | **16-20** |
| 07 | Visualization | **21-29** |
| 08 | Methodology for data preprocessing and feature extraction | **30-34** |
| 09 | Details about the predictive modeling process and performance metrics. | **35-38** |
| 10 | Insights on how named entities impact article engagement and popularity. | **39-41** |