## IJARSCT



International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 5, Issue 4, June 2025



## Characterization and Prediction of Popular Projects on GitHub

P. Srinivasa Rao<sup>1</sup>, Pathakala Lavanya<sup>2</sup>, Pochampally Shivani<sup>3</sup>, S S Nandini<sup>4</sup>, Kakkeri Venkat Sai<sup>5</sup>

Asst.Professor,ComputerScienceEngineering<sup>1</sup> Students, Computer Science Engineering<sup>2,3,4,5</sup> ACE Engineering College, Ghatkesar, India

**Abstract**: GitHub is a widely used platform for hosting open-source and private repositories, where the popularity of a repository is often determined by the number of stars it receives. This project, Characterization and Prediction of Popular Projects on GitHub, aims to analyze repository features and predict their popularity using machine learning techniques. The system extracts data from GitHub using its API, including key attributes such as stars, forks, watchers, issues, contributors, tags, programming language, and description. The collected data undergoes preprocessing, feature engineering, and model training using multiple regression algorithms, including Linear Regression, Random Forest, SGDRegressor, Ridge Regression, Lasso Regression, and Elastic Net Regression.

Keywords: GitHub, GitHub API, RESTAPI, Flask Web Application

Copyright to IJARSCT www.ijarsct.co.in



DOI: 10.48175/IJARSCT-27632



246