WEB PHISHING DETECTION

USING APPLIED DATA SCIENCE

TEAM ID : **PNT2022TMID44694**

TEAMMEMBERS : **DHIVYA S (732519104004)**

ABITHA J (732519104002)

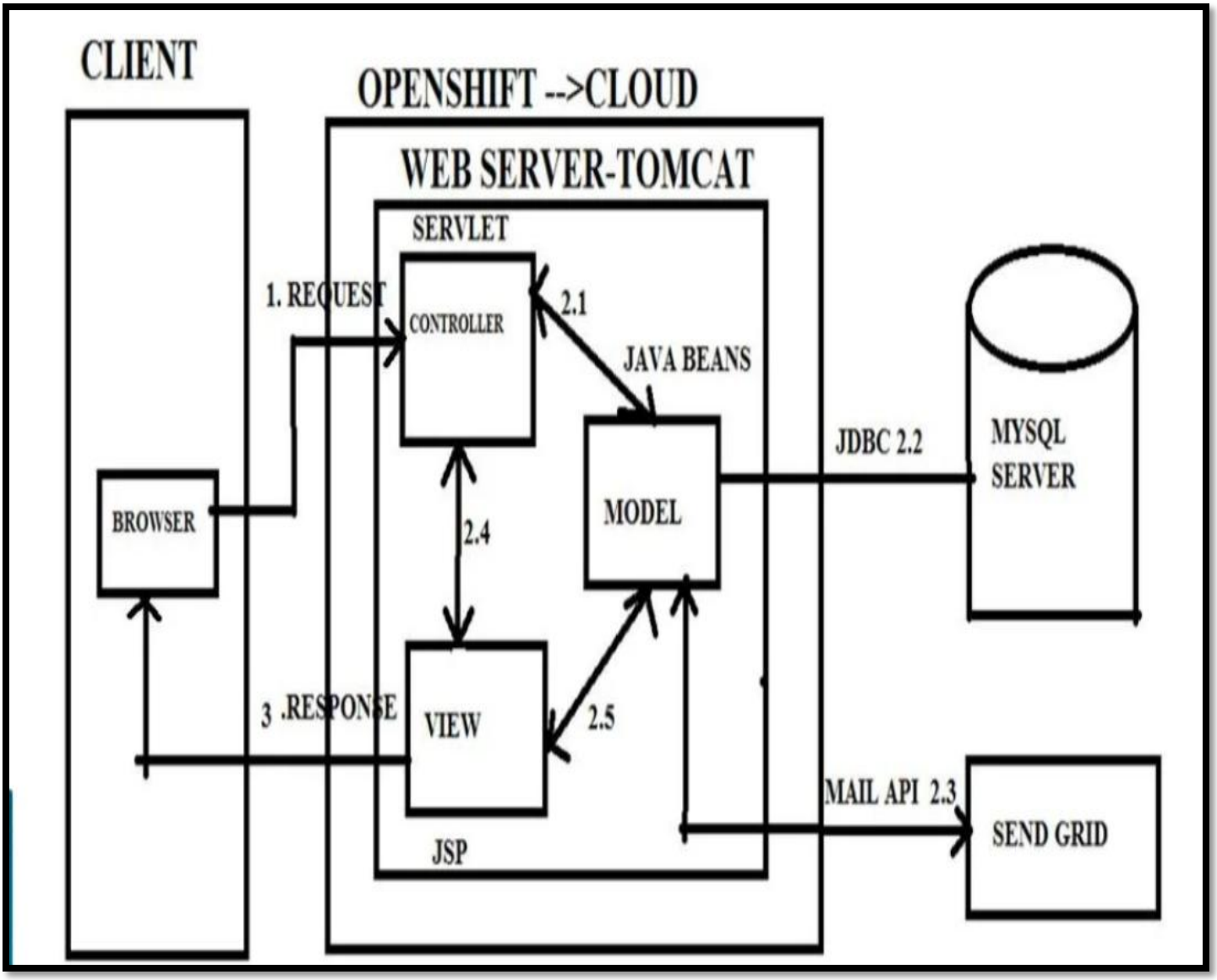
VASUNTHARA R (732519104030)

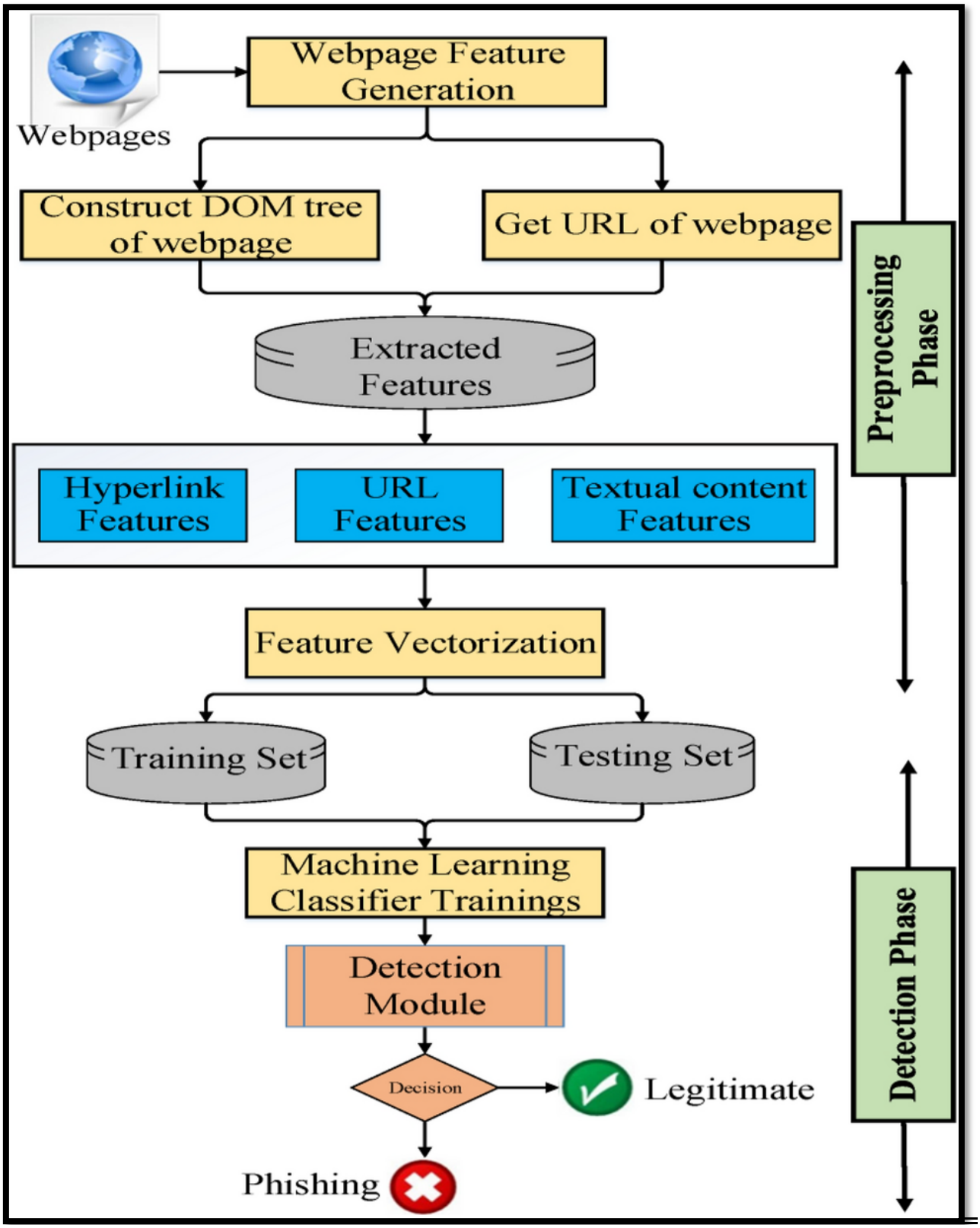
PRIYATHARSHINI G (732519104703)

DEPRTMENT : **CSE**

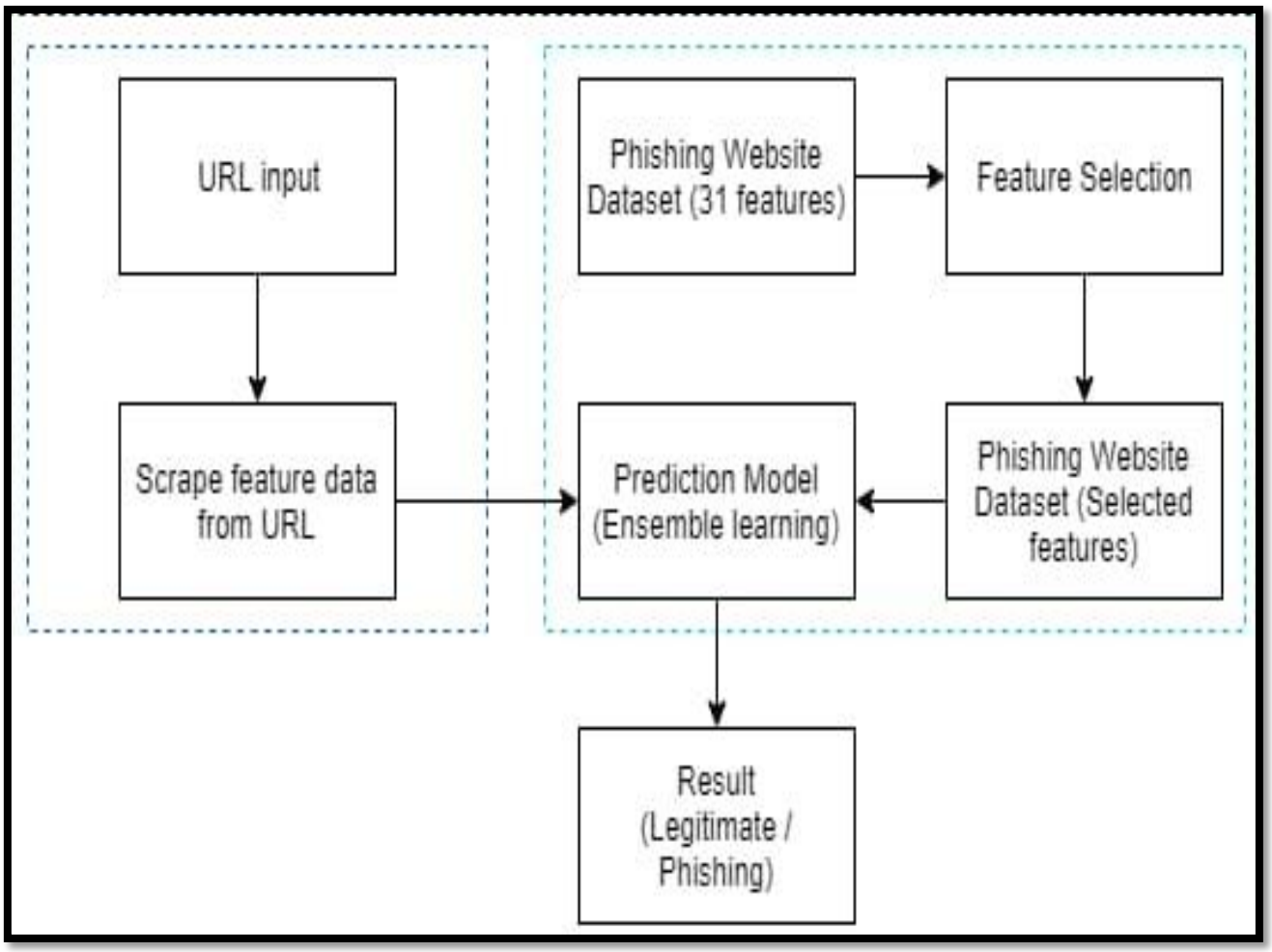
COLLEGE NAME : **SHREE VENKATESHWARA HI TECH ENGINEERING COLLEGE**

SOLUTION ARCHITECTURE




Webpages
Webpage Feature Generation
Construct DOM tree of webpage
Get URL of webpage
Extracted Features
Hyperlink Features
URL Features
Textual content Features
Feature Vectorization
Training Set
Testing Set
Machine Learning Classifier Trainings
Detection Module
Decision
Legitimate
Phishing
Preprocessing Phase
Detection Phase

FLOW CHART


URL input
Scrape feature data from URL
Phishing Website Dataset (31 features)
Feature Selection
Phishing Website Dataset (Selected features)
Prediction Model (Ensemble learning)
Result (Legitimate / Phishing)