

# San Francisco Crime Classification

Anthony Alves, Ian LaFlamme, Hunter Mack, Melisa Villmow

Faculty Advisor: Dr. Philip Chan, Dept. of Computer Sciences, Florida Institute of Technology

## Introduction

- Machine learning competition hosted by Kaggle.com
- 12 years of San Francisco arrests data given
- Goal is to predict the category of crime, given time and location
- Assist in building a safer society
- Work in a highly competitive leaderboard to evolve our skills
- Gain valuable experience with Data Mining and Machine Learning techniques

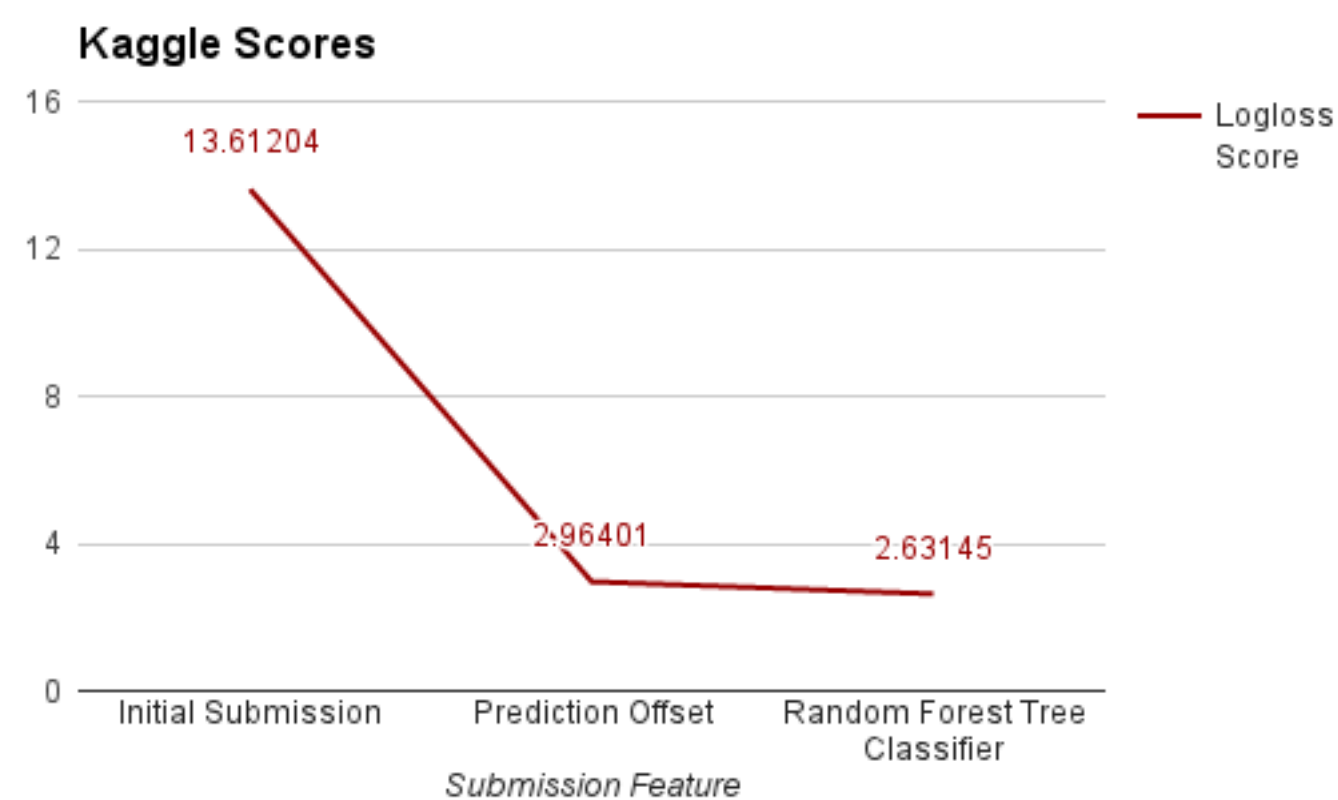


Figure 1: Submission Timeline

## Approach

- Utilize multiple Machine Learning libraries for more classifiers.
- Create a modular design for easy to maintain code
- Use an internal debugging module to improve our logloss score
- Derive additional attributes from the dataset such as hotspots and holidays

## Features

- Predict and Categorize Crime based off of past crimes
- Combines multiple algorithms from multiple languages to make the best predictions.
- Uses hotspots visualization to analyze crimes

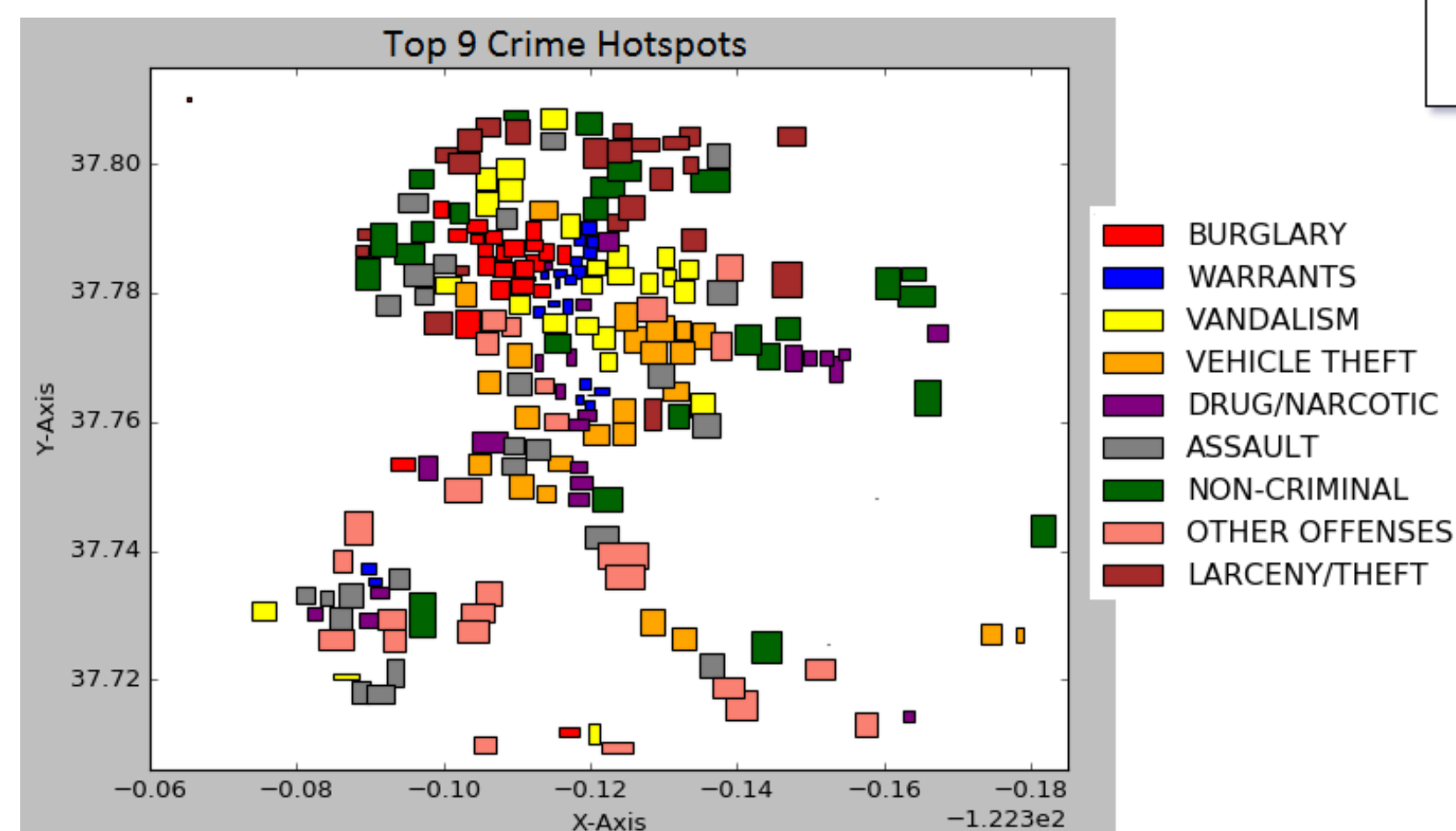


Figure 2: Hotspots Overlay

## System Architecture

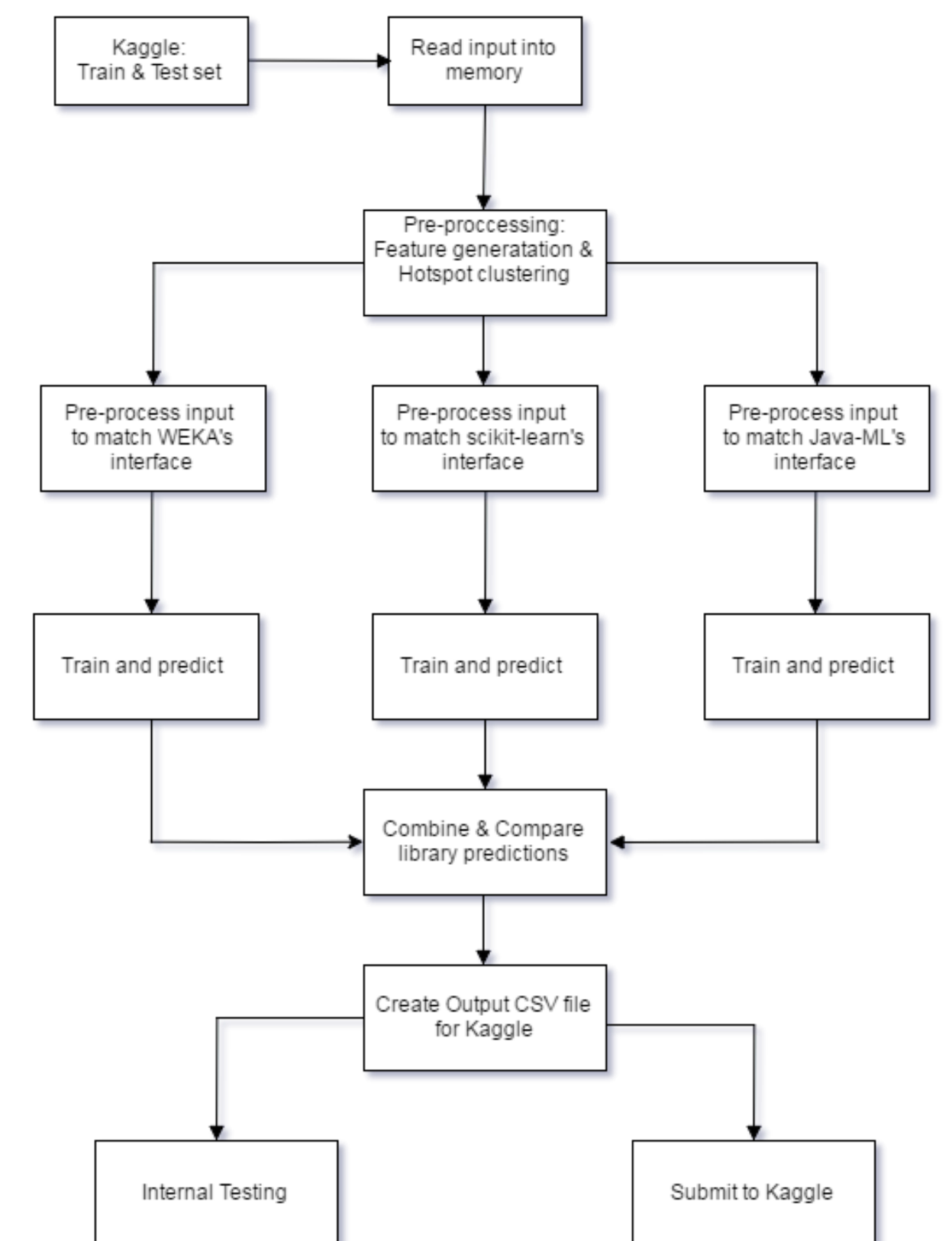


Figure 3: System Design

**NORTHROP GRUMMAN**



Engineering & Science  
Student Design Showcase  
at Florida Institute of Technology

