Plagiarism Detector

Team 207
Praveen Kumar Singh,
Parshva Shah,
Namrata Bilurkar

Motivation

It is a sad reality of college, a lot of students, intentionally or in desperation, copy off of other students' submissions or fail to cite references and violate the honor code. Objective of this project is to detect such cases of plagiarism among students.

Project Description

A web application developed as a solution to detect plagiarism among submissions to any given assignment in a course by students enrolled in that course at an institute.

Challenges

1. None of the team members had any experience with front-end development, which was a significant challenge.
2. Team size with only 3 members was the second challenge.

Features Implemented

1. Multiple strategies on demand for plagiarism detection.
2. Cron job is scheduled to run the detection system every night and sends a report to course staff.
3. Comparison of all submissions for plagiarism detection on demand.

Features Implemented cntd..

4. Similar lines detected in submissions are highlighted.
5. Support for multiple file submissions.
6. Usage statistics and system status is visible to the course staff.
7. System logs activity.
8. SMTP trap for error logs.

Strategies Implemented to detect plagiarism

1. Levenshtein distance on text
2. Longest common subsequence on text
3. Longest common subsequence on Abstract Syntax Tree
4. Tree edit distance on Abstract Syntax Tree
5. MOSS (Measure Of Software Similarity)
6. Overall score using a weighted polynomial

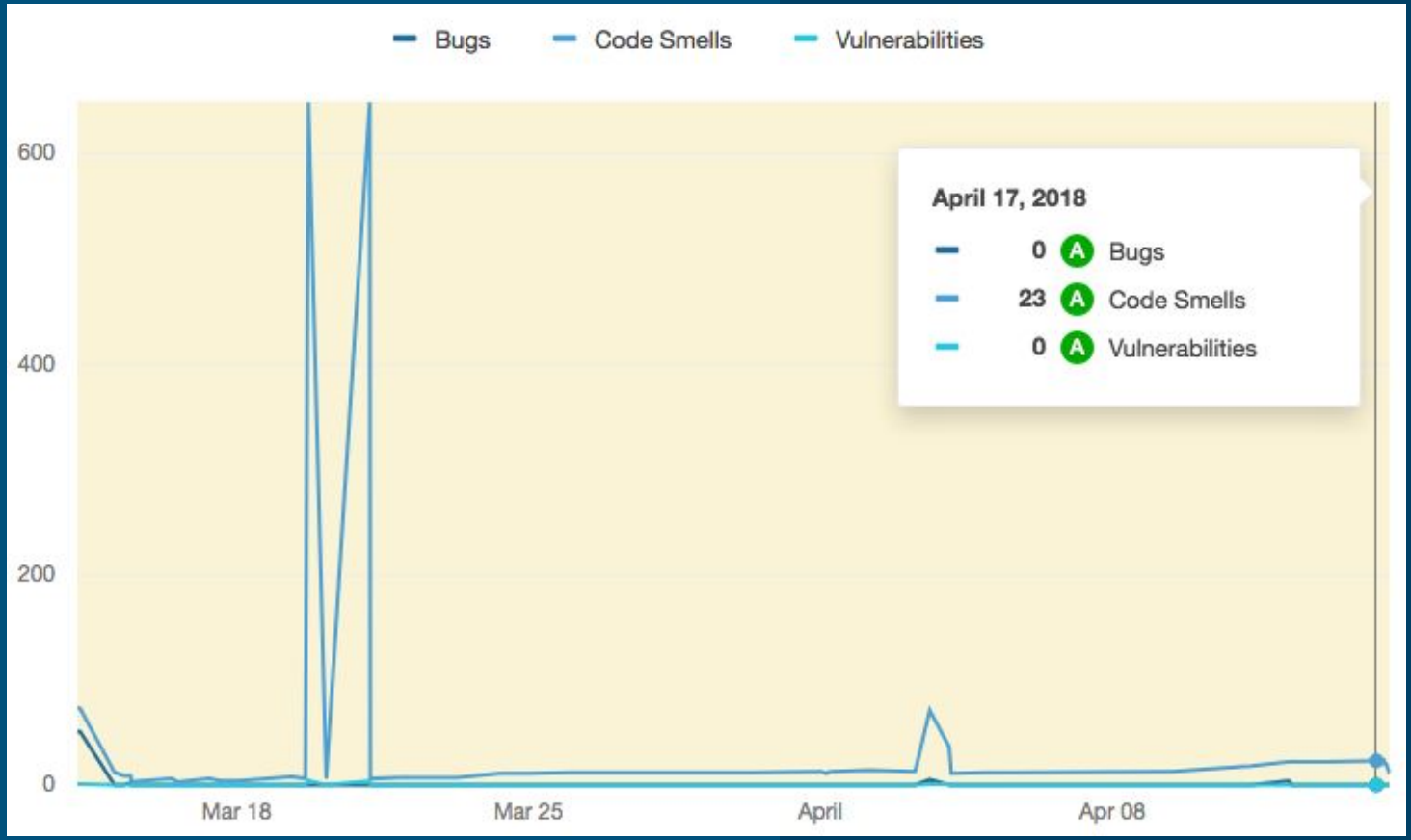
Process Followed

1. Entire Development was done in sprints of 2 weeks.
2. Smart commits were utilized to provide automation to for process.
3. Used slack for daily standups.
4. Continuous Integration was in place for each branch.
5. CI was automated for all the pull requests.
6. Utilized Slack for git and CI updates.

Process Followed cntd..

7. Used SonarLint in each developers system to ensure code quality.

8. Consistently maintained Grade on SonarQube over the entire course of the project.


Bugs
Code Smells
Vulnerabilities
600
400
200
0
Mar 18
Mar 25
April
Apr 08
April 17, 2018
0 A Bugs
23 A Code Smells
0 A Vulnerabilities

Tools used

1. Github
2. Jenkins
3. Jira
4. AWS EC2
5. AWS RDS
6. SonarQube
7. Slack

System stack

1. Java 8
2. SpringBoot
3. Swagger-UI v2.4
4. Log4j
5. Angular v1.6
6. Materialize CSS
7. JUnit
8. Maven

Future Scope

The Product is ready to be shipped and used for finding plagiarism, however it has scope for future development.

1. Social login and registration.
2. System support for detection in multiple programming languages.
3. Session management.
4. S3 storage for files and reports generated by the detection system.

Thank you