# Project Planning Phase
## Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)

| Date | 1 NOVEMBER 2022 |
|---|---|
| Team ID | PNT2022TMID27012 |
| Project Name | Web Phishing Detection |
| Maximum Marks | 8 Marks |

**Product Backlog, Sprint Schedule, and Estimation (4 Marks)**

Use the below template to create product backlog and sprint schedule

| Sprint | Functional Requirement (Epic) | User Story Number | User Story / Task | Story Points | Priority | Team Members |
|---|---|---|---|---|---|---|
| Sprint-1 | URL detector | USN-1 | URL is the first thing to analyse a website to decide whether it is a phishing or not | 10 | High | Girisha M V PITCHMA PRIYA K |
| Sprint-1 | | USN-2 | Some of URL-Based Features are<br>• Digit count in the URL<br>• Total length of URL<br>• Checking whether the URL is typo-squatted or not<br>• Checking whether it includes a legitimate brand name or not<br>• Number of subdomains in URL<br>• TLD is one of the commonly used one | 10 | High | Girisha M V PITCHMA PRIYA K |

| Sprint-2 | Domain detection | USN-3 | The purpose of Phishing Domain Detection is detecting phishing domain names. Therefore, passive queries related to the domain name, which we want to classify as phishing or not, provide useful information to us. | 10 | High | SHRUTHI R SURIYA LEKSHMI R M GIRISHA M V |
|---|---|---|---|---|---|---|
| Sprint-2 | | USN-4 | Some useful Domain-Based Features are<br><br>• Its domain name or its IP address in blacklists of well-known reputation services?<br><br>• How many days passed since the domain was registered?<br><br>• Is the registrant name hidden? | 10 | High | SHRUTHI R SURIYA LEKSHMI R M GIRISHA M V |
| Sprint-3 | Page based features and Content based features | USN-5 | Page-Based Features are using information about pages which are calculated reputation ranking services.<br>Obtaining these types of features requires active scan to target domain. Page contents are processed for us to detect whether target domain is used for phishing or not | 10 | High | Girisha M V PITCHMA PRIYA K |

| | | | | | | |
|---|---|---|---|---|---|---|
| Sprint-3 | | | • Global pagerank<br>• Country pagerank<br>• Position at the Alexa top 1 million site<br>Some processed information about pages are<br>• Page titles<br>• Meta tags<br>• Hidden text<br>• Text in the body<br>• Images etc. | | | Girisha M V<br>PITCHMA PRIYA K |
| Sprint-4 | Detection process | USN-6 | Detecting Phishing Domains is a classification<br><br>problem, so it means we need labeled data<br><br>which has samples as phish domains and legitimate domains in the training phase | 20 | | PITCHMA PRIYA K<br><br>SURIYA LEKSHMI R M<br><br>SHRUTHI R<br><br>GIRISHA M V |

**Project Tracker, Velocity & Burndown Chart: (4 Marks)**

| Sprint | Total Story Points | Duration | Sprint Start Date | Sprint End Date (Planned) | Story Points Completed (as on Planned End Date) | Sprint Release Date (Actual) |
|---|---|---|---|---|---|---|
| Sprint-1 | 20 | 6 Days | 24 Oct 2022 | 29 Oct 2022 | 10 | 29 Oct 2022 |
| Sprint-2 | 20 | 6 Days | 31 Oct 2022 | 05 Nov 2022 | 10 | 05 Nov 2022 |
| Sprint-3 | 20 | 6 Days | 07 Nov 2022 | 12 Nov 2022 | 10 | 12 Nov 2022 |
| Sprint-4 | 20 | 6 Days | 14 Nov 2022 | 19 Nov 2022 | 20 | 19 Nov 2022 |

**Velocity:**

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

**Burndown Chart:**

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.