Assignment 1-Part B

SENG8100 – Agile Software Prototyping

Pradeepti Kasam- 8965985

TASK B:

Task B.1 :

Demonstrate the Product Vision using a physical scrum board (cardboard.it). Record your observations/reflections to review the importance of features considered when building the proposed product. (Including a screenshot of your physical scrum board is mandatory).

Product Vision Board :

VISION To make text annotation seamless and ensure efficient retrieval for multilingual data in regional languages.			
TARGET GROUPS	**NEEDS**	**PRODUCT**	**BUSINESS GOALS**
Annotators: Individuals tasked with labeling and tagging text content in multiple languages. **Researchers & Linguists:** Focusing on analyzing and processing textual data in non-English or regional languages. **Business Analysts:** Companies needing effective tools for text annotation and sentiment analysis to improve customer engagement.	Ability to annotate texts in multiple languages. Automatic entity tagging (Named Entity Recognition). Saving and tracking annotation progress efficiently. Integration of machine learning for suggestions and disambiguation. Robust text retrieval for better searchability in regional languages.	**Text Annotation Interface:** • Highlight and label text segments. • Named Entity Recognition (NER) for tagging. • Autosave features for annotation progress. **Data Storage:** • Secure storage of annotated data. • Retrieval system optimized for multilingual search algorithms. **Language Processing:** • Contextual word disambiguation. • Machine-learning recommendations for annotators.	• Subscription-based plans tailored for different users (Annotators, Researchers, Businesses). • Free version with limited features for small-scale users. • Paid tiers: Basic, Professional, and Enterprise levels offering advanced features.

Fig 1.1 : Product Vision Board

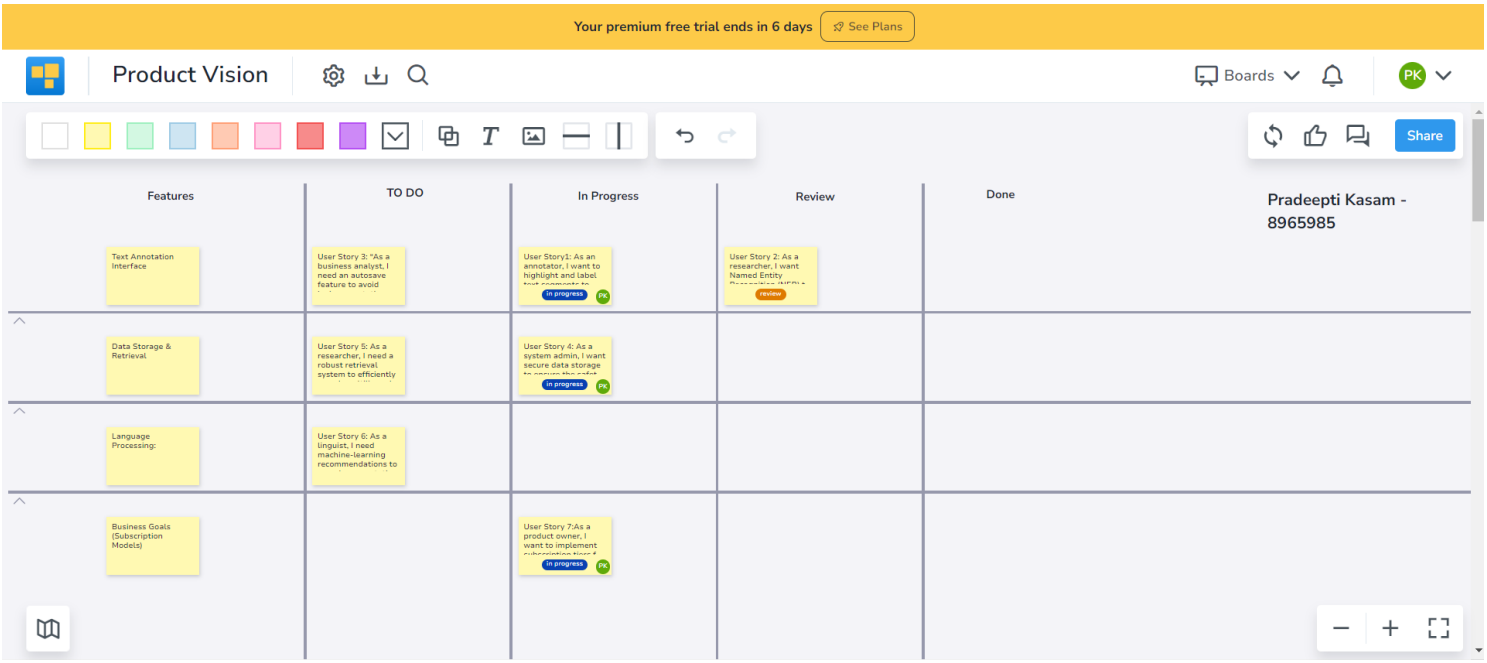

Fig 1.2 : Scrum Board for Product Vision

Review Reflections/Observations:

- **Feature Prioritization:** Some features (like the Text Annotation Interface) are important to the idea and would likely be highlighted in the early stages. Dependencies and bottlenecks can be identified by watching how fast and effectively they can flow across the board.
- **Business Model:** Evaluating how well the product is in line with business goals is made easier by keeping an eye on its positioning and development of features associated with the subscription model.
- **Multilingual Processing:** It will probably be noted that the NER labeling and complexity of machine-learning recommendations require additional time. Monitoring their progress will demonstrate how important it is to provide precise language tools.

Task B.2 :

Discuss the user persona that you are considering for building levels of activity/epic and user stories. Record your observations/reflections to review the importance of features considered when building your product. (Including a screenshot of your physical scrum board is mandatory)

User Personas

1. **Annotator (Language Expert)**:
    - **Background**: Language experts, researchers, or professionals working in linguistics. They have a deep understanding of regional languages and are responsible for annotating and tagging the text.
    - **Goals**: Ensure accurate and contextually relevant annotations. Make meaningful contributions to the growing data of regional languages.

2. **Admin (Project Manager/Lead Researcher)**:
    - **Background**: Managers overseeing multiple annotation projects. They are responsible for dataset management, user roles, and ensuring the project's overall success.
    - **Goals**: Assign tasks to annotators, manage datasets, review and approve annotations, ensure deadlines are met.
3. **Researcher/Developer**:
    - **Background**: NLP or AI researchers using the annotated data to train machine learning models or conduct research in language processing.
    - **Goals**: Access high-quality, labeled data for regional languages, integrate annotations into their ML pipelines.
4. **End-User (Knowledge Seeker)**:
    - **Background**: General users or professionals who want to retrieve information from the annotated content, such as educators, translators, or students.
    - **Goals**: Retrieve accurate and contextually relevant information from regional language texts.

Epic-Level User Stories

1. **Annotator Role (Language Expert):**
    - Epic: As an annotator, I want to annotate and label text content in regional languages to ensure that the data is accurate and contextually relevant.
        - User Story 1: As an annotator, I want to highlight and label entities (people, places, objects) in regional language text.
        - User Story 2: As an annotator, I want my annotations to be saved to the database for future review.
        - User Story 3: As an annotator, I want to create custom labels for entities in the dataset for more flexible annotations.
2. **Admin Role (Project Manager):**
    - Epic: As an admin, I want to manage annotators, datasets, and workflows to ensure the smooth progress of the annotation project.
        - User Story 1: As an admin, I want to upload datasets for annotators to work on.
        - User Story 2: As an admin, I want to assign datasets to specific annotators.
        - User Story 3: As an admin, I want to review annotations submitted by annotators to ensure quality and accuracy.

3. **Researcher Role:**
    - Epic: As a researcher, I want to access the annotated datasets to use them in machine learning models and language studies.
        - User Story 1: As a researcher, I want to download the annotated dataset in CSV or JSON format, so I can use it in my NLP pipeline.
        - User Story 2: As a researcher, I want to retrieve annotation statistics (like frequency of certain tags), so I can understand the dataset's structure.
4. **End-User Role (Knowledge Seeker):**
    - Epic: As an end-user, I want to search and retrieve text information from regional languages based on my query to access the relevant content.
        - User Story 1: As an end-user, I want to search for a keyword and see the relevant text segments, so I can gather the information I need.
        - User Story 2: As an end-user, I want the system to suggest similar content in other regional languages, so I can broaden my knowledge base.

Reflections and Observations:

1. **User-Centered Design:**

   It is essential that every aspect of the annotator interface remains simple. Tools should enable users to annotate meaningfully without facing technical challenges, especially in light of the complexities of regional languages.

2. **Annotation Flexibility:**

   In order for annotators to function with a variety of regional languages efficiently, custom labels must be allowed. Flexible methods are required to capture language-specific details, which standardized systems might not be able to support.Features that save time and guarantee data integrity are non-negotiable. This keeps long-term annotation efforts going and prevents data loss.

3. **Scalable Dataset Management:**

   To work with big datasets and oversee several projects and annotators at once, administrators require strong tools. Maintaining projects on time requires the capacity to redistribute work, monitor advancement, and guarantee that due dates are fulfilled.Administrators must have a strong quality control system in place to ensure that the data they create is accurate. Without it, there is a greater chance of low-quality annotations, which compromises the overall dependability of the output.

4. **Researcher-Driven Features:**

   Giving researchers access to well-structured, readily exportable data guarantees that they can incorporate the annotations into their machine learning models fast.By providing information such as annotation statistics, researchers may assess the worth of the dataset without having to manually explore it, which increases productivity and helps with decision-making.

5. **Powerful Search and Retrieval for End-Users:**

   For the end-user, accurate and contextually relevant search results are a priority. Regional language texts can be complex, and poor search functionality can render the data unusable for them.Suggesting similar content in other languages is a feature that sets the product apart. This fosters greater knowledge sharing and cultural exchange across language barriers, fulfilling a key societal role in broadening access to information.

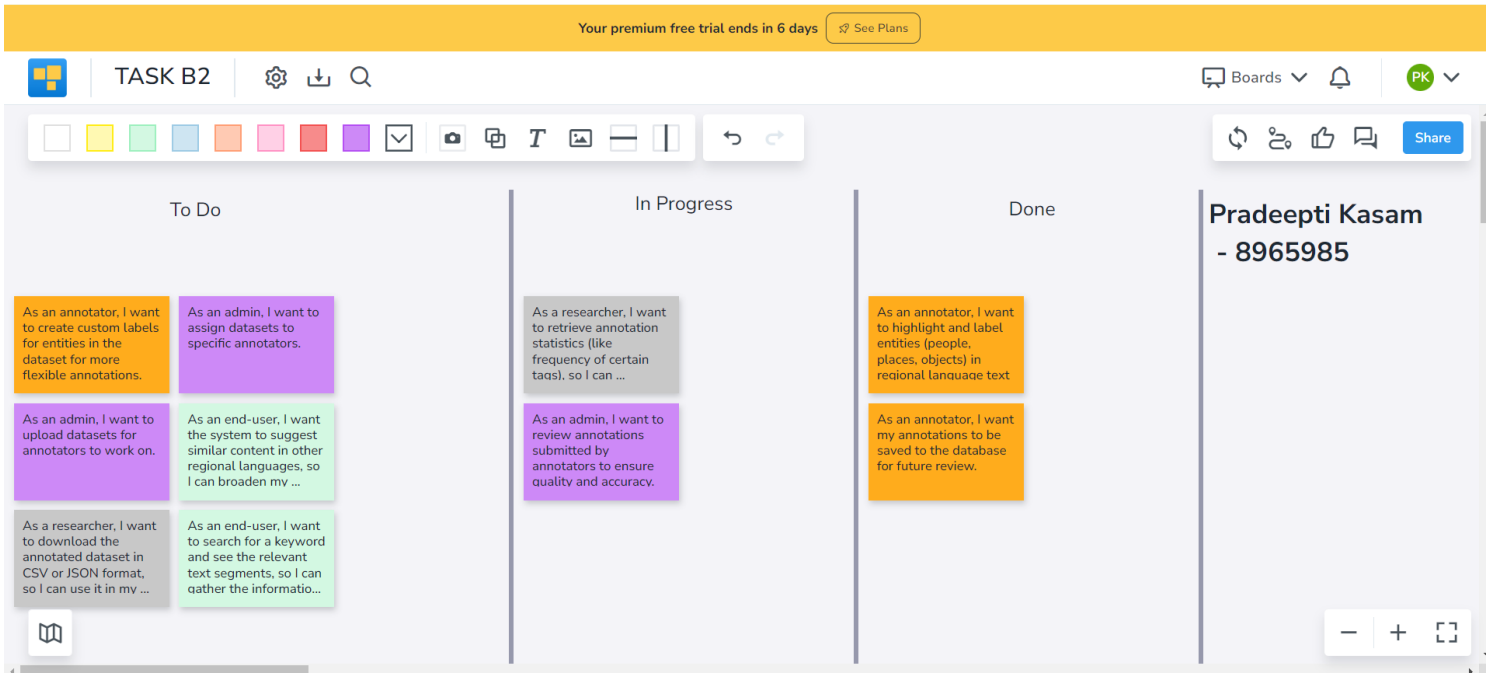

Fig2: Scrum board of user stories based on user persona

Task B3:

Create the backbone for the Product Vision by considering the high-level features and build the body with the levels of activity/epic and user stories. Record your observations/reflections to review the importance of features considered when building your product. (Including a screenshot of your physical scrum board is mandatory).

High-Level Features (Backbone):

1. Text Annotation Interface:

- **Epic:** Provide a user-friendly interface for annotators to interact with text data.
- **User Story:** As an annotator, I want to highlight and label specific parts of the text (e.g., words, sentences) for analysis.

2. Entity Tagging Functionality

- **Epic**: Allow tagging of various entities within the text (e.g., people, locations, dates).
- **User Story** : As an annotator, I want to tag entities such as persons, organizations, and dates to enrich the dataset.

3. Autosave for Annotations

- **Epic:** Automatically save progress to prevent data loss.
- **User Story:** As an annotator, I want the system to autosave my annotations, so I don't lose progress due to system errors.
- **User Story:** As an annotator, I want to be notified when changes are made to a dataset I'm working on, such as new content or modifications.

4. Sentiment Analysis Integration

- **Epic:** Implement sentiment analysis capabilities for dataset evaluation.
- **User Story:** As an admin, I want to integrate sentiment analysis for specific datasets to analyze emotional tone in text.

5. AI-Powered Annotation Suggestions

- **Epic:** Incorporate AI-based annotation suggestions for improved annotation efficiency.
- **User Story:** As an annotator, I want AI-based suggestions for possible annotations to speed up the labeling process.
- **User Story:** As an annotator, I want to dispute rejected annotations and provide explanations for my decisions.
- **User Story:** As an annotator, I want the system to suggest tags based on text content to speed up the annotation process.

6. Comment Sentiment Analysis

- **Epic:** Implement sentiment analysis for feedback and internal comments.
- **User Story:** As an admin, I want to analyze the sentiment of comments left by annotators to assess team morale or feedback.

7. AI-Powered Dataset Insights

- **Epic:** Enable AI-driven insights into datasets for enhanced data understanding.
- **User Story:** As an admin, I want AI to analyze datasets and provide insights, such as commonly tagged entities or difficult sections.

8. Download Annotated Dataset

- **Epic:** Allow dataset export for use in external analysis or pipelines.
- **User Story:** As a researcher, I want to download the annotated dataset in CSV or JSON format, so I can use it in my NLP pipeline.

Observations and Reflections

Core Features (Text Annotation, Entity Tagging, Autosave): These are essential to how the product works. They ensure the basic usability and usefulness of the platform, notably for annotators.

AI and ML Features (Sentiment Analysis, AI-Powered Suggestions, AI-Powered Insights): These are very valuable since they increase productivity, automate tasks, and provide more in-depth information. In highly competitive markets, they are important differentiators that can help the platform stand out.

Utility and Team Support Features (Download Annotated Dataset, Comment Sentiment Analysis): These capabilities improve the product's usefulness for both internal (team management) and external (data exporting) operations.

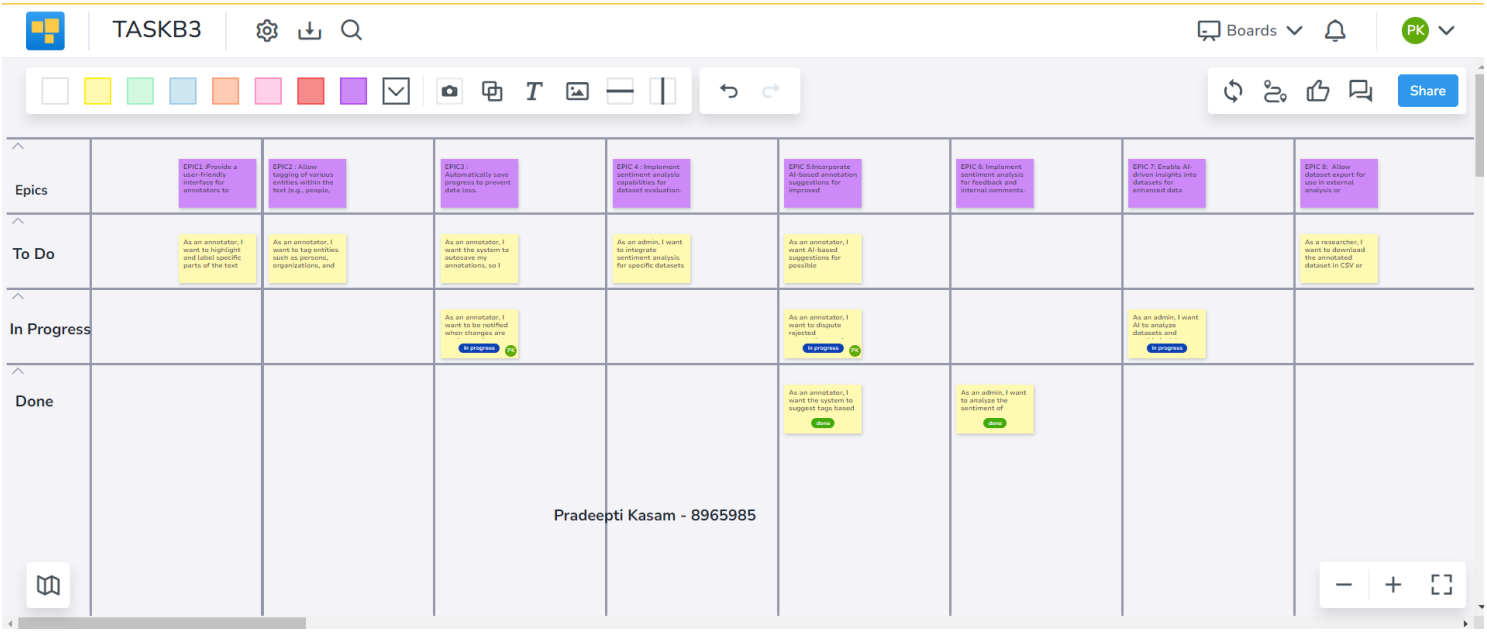

Fig3: Scrum board for user stories based on High-Level Features

Task B.4:

Analyze the user stories and slice out the minimum viable product features to satisfy the user persona that has been considered. You can consider using horizontal lines plus persona cards (cards with user images). Record your observations/reflections to review the importance of features when building your product. (Including a screenshot of your physical scrum board is mandatory).

User : Annotator

Key MVP Features:

1. **Text Annotation Interface :**
   - User Story : As an annotator, I want to highlight and label entities (people, places, objects) in regional language text.

2. **Entity Tagging Functionality:**
   - User Story : As an annotator, I want to tag entities such as people, places, and objects.

3. **Annotation Data Storage:**
   - User Story : As an annotator, I want my annotations to be saved to the database for future review.

4. **Autosave for Annotations:**
   - User Story : As an annotator, I want the system to autosave my annotations so I don't lose progress.

5. **Annotation Progress Tracking:**
   - User Story : As an annotator, I want to see my progress for the dataset I'm annotating.

User :Admin

1. **Dataset Upload :**
   - User Story : As an admin, I want to upload datasets for annotators to work on.

2. **Dataset Assignment:**
   - User Story : As an admin, I want to assign datasets to specific annotators.

3. **Review Interface :**
   - User Story : As an admin, I want to review annotations submitted by annotators to ensure quality and accuracy.

4. **Accept/Reject Annotations:**
   - User Story : As an admin, I want to accept or reject annotations made by annotators.

5. **RBAC for Annotators/Admins:**
   - User Story : As an admin, I want to manage roles and permissions to control who has access to different parts of the system.

User : Researcher

1. **Download Annotated Dataset:**
   - User Story : As a researcher, I want to download the annotated dataset in CSV or JSON format, so I can use it in my NLP pipeline.

2. **Retrieve Annotation Statistics:**
   - User Story : As a researcher, I want to retrieve annotation statistics (like frequency of certain tags), so I can understand the dataset's structure.

3. **API Access for External Systems:**
   - User Story : As an admin, I want external systems to access annotations via an API to integrate with other platforms.

User : End-User

1. **Keyword Search for Relevant Segments:**
   - User Story : As an end-user, I want to search for a keyword and see the relevant text segments, so I can gather the information I need.

2. **Suggest Similar Content in Regional Languages:**
   - User Story : As an end-user, I want the system to suggest similar content in other regional languages, so I can broaden my knowledge base.

Observations and Reflections

1. **Prioritizing Core Functionalities:** The MVP focuses on key components that allow for basic functionalities like as dataset administration, data retrieval, and annotation. This makes sure the basic needs of the user are satisfied without making the first release unduly complicated.
2. **Quick Feedback Loop:** These MVP features enable the product to rapidly collect user feedback, which is essential for iterative development and enhancement.
3. **Scalability and Flexibility:** Based on user input and changing requirements, the selected features offer a framework that can be readily expanded with other functionality, such AI-powered suggestions or advanced analytics.
4. **User-Centric Design:** The essential needs of each character are met, guaranteeing that the product will be beneficial to all parties involved.

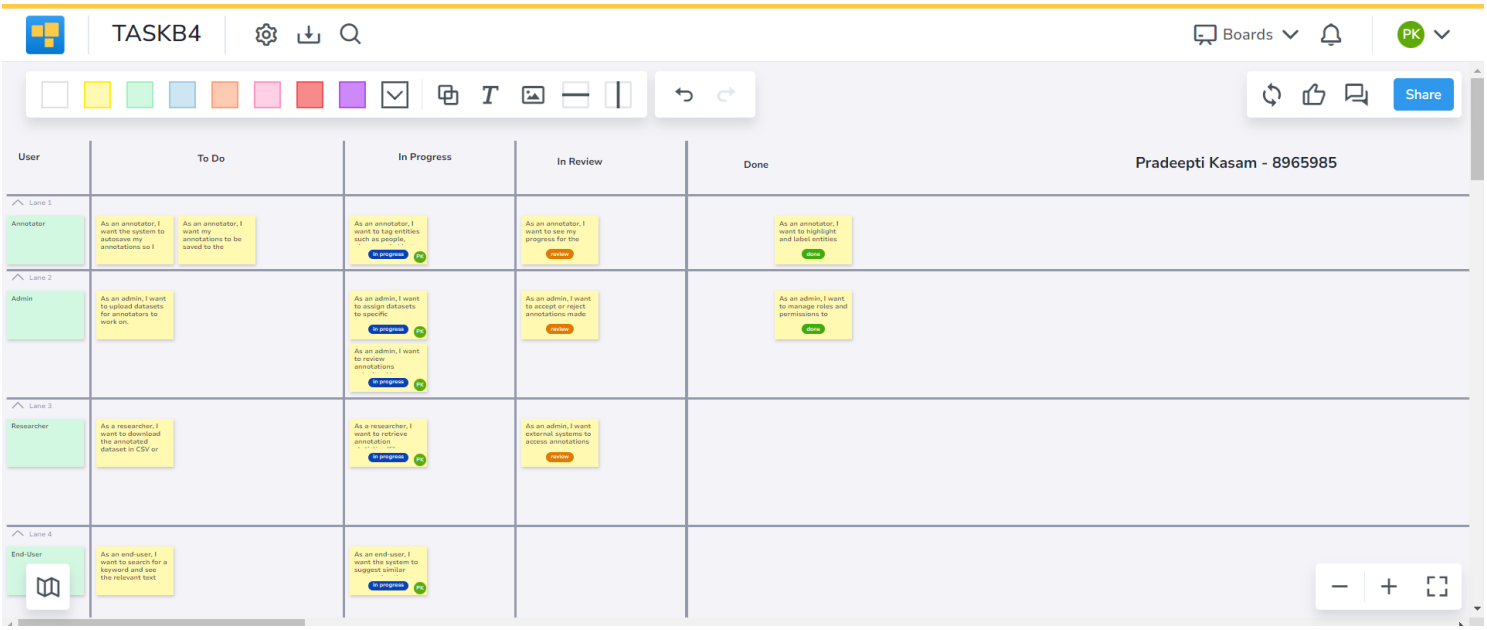

Fig4: Scrum board for user stories based on Minimum Viable Product

Task B.5:

Create the whole product backlog from the user story mapping by considering the key activities.

ProductBacklog - Assignment1-Pradeep

Task B.6: Create the above story mapping and provide answers to the following questions:

1. **What is the importance of using a user persona in building the user story mapping?**

   When developing user story mapping, the team can concentrate on delivering essential and relevant functionality by using a user persona to ensure that features are customized to the unique demands and behaviors of actual users.

2. **How do you think user story mapping helps you to achieve the triad of feasibility, engineering, and usability especially in the domain of ML application that you proposed?**
   - By aligning the features with the data, technology, and resources at hand and decomposing complex ML workflows into manageable sections, it guarantees that the features are practical.
   - It helps prioritize core engineering tasks (like data processing, model training) while maintaining focus on user experience (such as easy annotation tools for annotators), ensuring the system is functional, scalable, and user-friendly.

3. **What is the significance of minimum viable product release, especially in the domain of the ML applications already available?**

   The significance of a minimum viable product (MVP) release in ML applications is:

   Early Feedback: Before extensively investing in complex features, the team can utilize it to test fundamental features rapidly, get real-world customer feedback, and refine the model or functionality.

   Reducing Risk: An MVP helps distinguish the product early in an industry with a lot of machine learning applications by verifying assumptions and making sure the solution effectively addresses user concerns without investing money on unnecessary features.

References :

Harish, B. S., & Rangan, R. K. (2020). A comprehensive survey on Indian regional language processing. In *SN Applied Sciences* (Vol. 2, Issue 7). Springer Nature. https://doi.org/10.1007/s42452-020-2983-x

Schwaber, K., & Sutherland, J. (2020). *The Scrum guide: The definitive guide to Scrum: The rules of the game*. Scrum.org. https://www.scrumguides.org/scrum-guide.html