A blue and black logo

Description automatically generated

**Mansoura University, Egypt**

Faculty of Computer & Information Sciences

**Computer Science Department**



**Vice Dean for Education and Student Affairs of Faculty of Computer & Information Sciences, Mansoura University**

We will be delighted if you accept our application for our graduation project to participate in the “Dell Envision The Future” Graduation projects competition

Dell Technologies has launched the annual graduation project competition for senior undergraduate students from universities in the Middle East, Russia, Africa, and Turkey.

The competition is intended to spark the creativity of students for their graduation projects to play an active role in the transformation of IT and get the opportunity to shine and win prizes.

**Table of Content**

1. **Team Members & Faculty Advisor**
2. **Project**
   1. Title
   2. Which sector does your project tackle?
   3. Which technology area are you using to tackle the addressed problem?
   4. Question 1: What is the main problem that you are solving? Why did you choose it?
   5. Question 2: Give a brief overview of the current solutions that you found in your literature survey.
   6. Question 3: Describe how your proposed solution will be used and operated in a real-life environment?
   7. Question 4: Describe the high-Level architecture and the main software and hardware components of your system. (You can use block diagrams)
   8. Question 5: What are the technology platforms that you intend to use in building your system, such as operating systems, programming languages, backend/frontend stacks, AI models, etc.?
   9. Question 6: Compared to the existing solutions that you have read about, what are the new ideas that you intend to develop and include in your solution?
   10. Question 7: Describe your teams system development methodology and quality assurance process.
   11. Question 8: Give your initial project management plans, especially: distribution of responsibilities and tasks, milestone schedule, training plan, risk management, contingency plans, and change control procedures.
3. **Team Members**

|  |  |  |
| --- | --- | --- |
| **Name** | **Email** | **Gender** |
| Amr Mohamed Shoukry | shoukryworkamr1@gmail.com | M |
| Nouran Adel Elkhouly | nouran.elkhouly12@gmail.com | F |
| Saad Al-Tohamy Saad | saadaltohamy@std.mans.edu.eg | M |
| Shahd Faris El-Badrawy | shahdfaris1623@gmail.com | F |
| Abdlerahman Karim Mossad | abdelrahmankarimwork@gmail.com | M |
| Salma Ali El-Sherbiny | salmaegy10@gmail.com | F |

**Faculty Advisor**

**Name:** Sara El-Sayed Youssef El-Metwally

**Email**: [sarah\_almetwally4@mans.edu.eg](mailto:sarah_almetwally4@mans.edu.eg)

**University Job Title:** Assistant Professor in the Department of Computer Science

1. **Project**
2. **Title**- ArchivAI
3. **Which sector does your project tackle?  
   -** Wellbeing
4. **Which technology area are you using to tackle the addressed problem?**- AI
5. **Question 1: What is the main problem that you are solving? Why did you choose it?**  
   - Managing documents manually has become complex and difficult with the increasing volume of files that organizations deal with. Employees spend a lot of time searching and organizing documents, leading to delays in making decisions and increasing the possibility of losing crucial documents. Traditional document management systems do not keep pace with the increase in file size and do not meet changing business needs, causing compliance and security issues.

We chose this problem because we really feel it every day in our lives, dealing with a large amount of documents, organizing them and searching for them is very tiring and also a waste of time.

Especially when you are forced to organize thousands and maybe millions of documents in large companies so that nothing important is lost in the future, so we wanted to solve this problem and make it easy for people and companies to organize this huge number of documents and also search for them with ease

1. **Question 2: Give a brief overview of the current solutions that you found in your literature survey.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Feature | Describe of feature | DOKMEE | DocuWare | M-Files | ArchivAI |
| Folder Creation and Initial Setup | The system lets users manually organize files, creating a structure for AI to learn from. | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Don’t Have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill |
| Automatic File Categorization with AI | After setup, the AI automatically sorts new files based on content, eliminating manual organization. | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |
| Semantic Search for Easy Document Retrieval | We will implement semantic search to allow users to find documents by meaning, enhancing search speed and accuracy. | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill |
| Text Extraction with OCR | The system will use OCR to extract text from scanned documents, enabling searches for non-digital content. | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |
| File Parsing and Metadata for Better Search | Each document will be scanned for key details, saved as metadata, and used to simplify searches. | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |
| Testing and Feedback | We will test the AI for accurate categorization and smooth searches, using user feedback to improve the system. | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill |
| Deployment and Ongoing Updates | After testing, the system will be deployed for users, with continuous updates based on feedback and technological advancements. | Have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |

1. **Question 3: Describe how your proposed solution will be used and operated in a real-life environment?**- Real-Life Application Scenario

In this scenario, the user will define the file structure and classifications, and place sample documents into each folder. This will allow the AI model to train on these documents.  
  
1) Defining Structure:

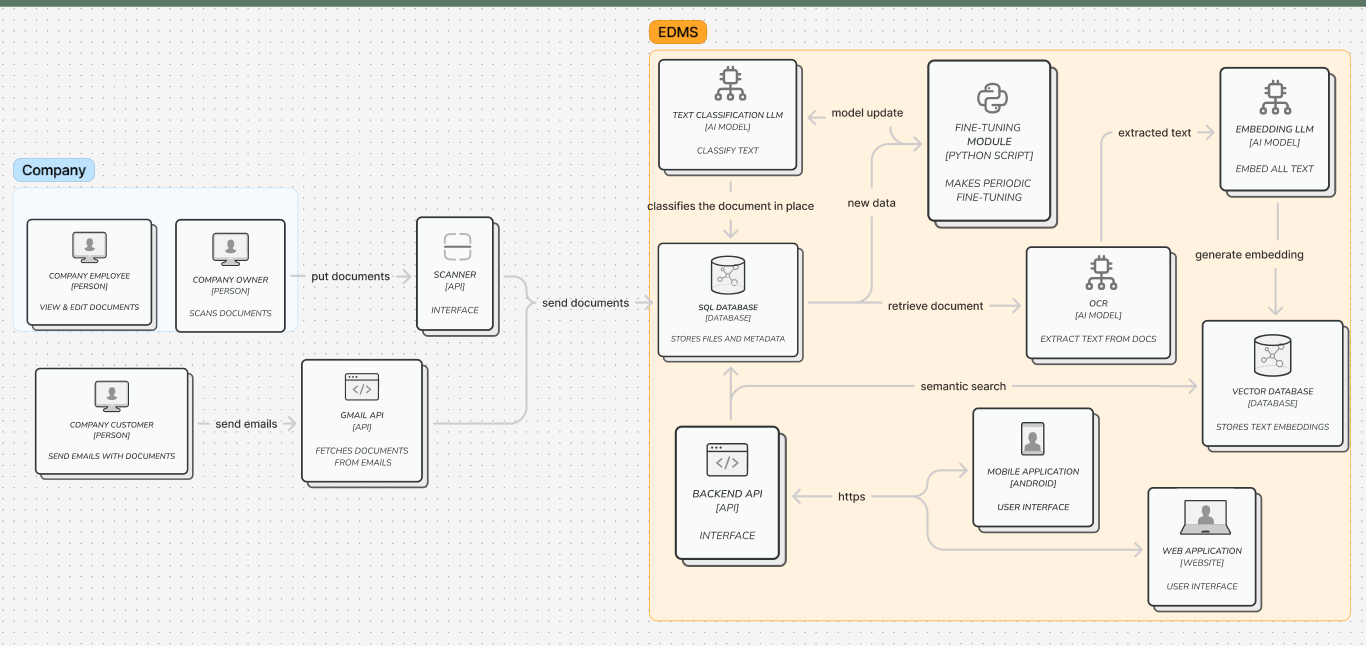
The user creates a specific file structure and assigns appropriate classifications for each type of document.  
  
2) Setting Up Sample Documents:

The user places samples of documents in each folder, helping the AI learn how to categorize new documents based on this training data.  
  
3) Uploading New Documents:

When the user uploads new documents, the AI will analyze their content and determine the  storage location based on  previous training and classifications.

4) Semantic Search Feature:

When the user wants to find a specific document, they can utilize the semantic search feature to locate documents based on their content or What did he remember about this document.

1. **Question 4: Describe the high-Level architecture and the main software and hardware components of your system. (You can use block diagrams)**  
     
   
2. **Question 5: What are the technology platforms that you intend to use in building your system, such as operating systems, programming languages, backend/frontend stacks, AI models, etc.?**-

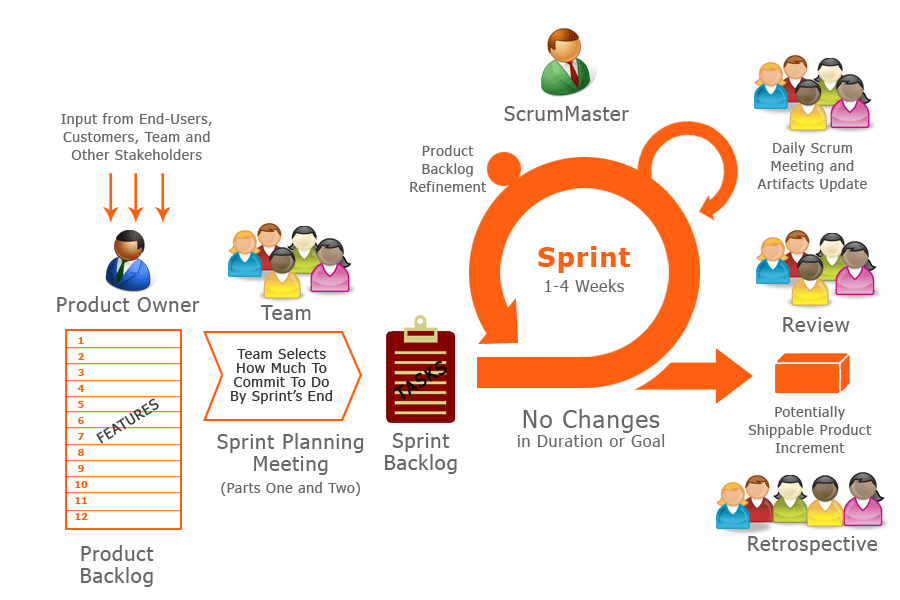
|  |  |  |  |
| --- | --- | --- | --- |
| **Purpose** | **Operating System** | **Technology stack** | **Programming Language** |
| **Mobile Application** | **Android** | **Android Studio** | **Kotlin** |
| **Web Application** | **Cross Platform** | **Next.js** | **Typescript** |
| **Backend** | **Cross Platform** | **ASP.NET** | **C#** |
| **AI Models** | **Cross Platform** | **PyTorch** | **Python** |

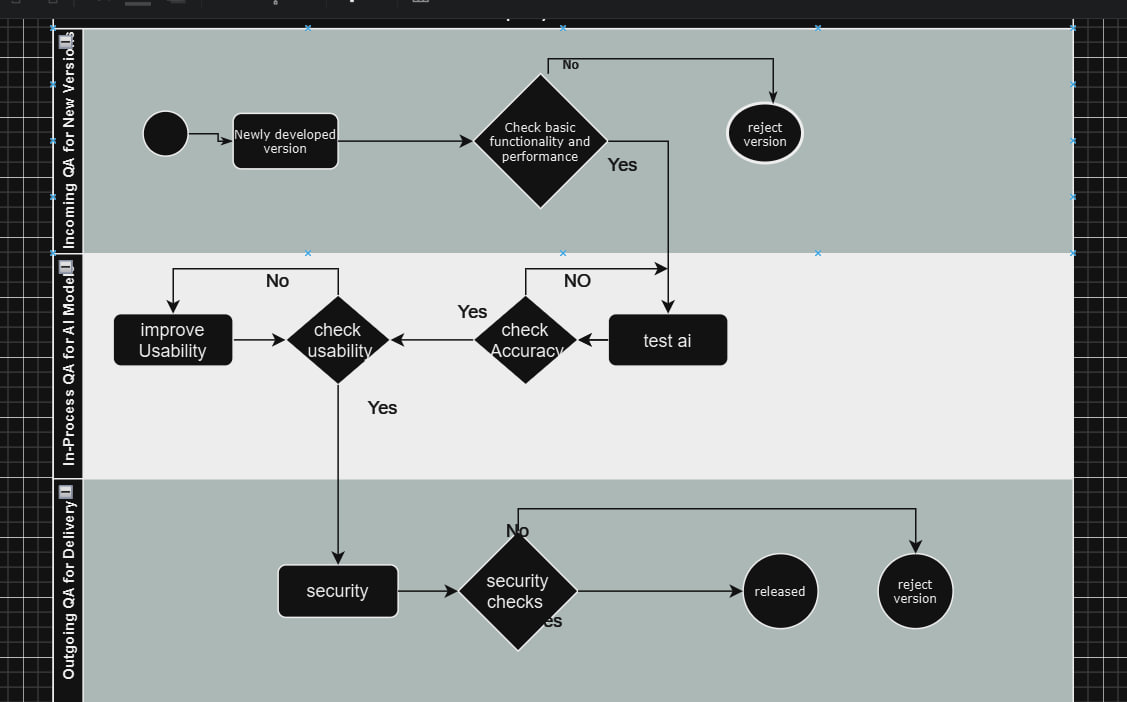
|  |  |  |
| --- | --- | --- |
| **Component Type** | **Component Name** | **Description** |
| **Software** | **OCR (Optical Character Recognition) [AI Model]** | **Extracts text from documents** |
| **Software** | **Text Classification LLM [AI Model]** | **Classifies the text content of the documents** |
| **Software** | **Embedding LLM [AI Model]** | **Embeds text into vector representations for further processing** |
| **Software** | **Fine-Tuning Module [Python Script]** | **Periodically fine-tunes the text classification model with new data** |
| **Software** | **SQL Database [Database]** | **Stores the document files and related metadata** |
| **Software** | **Vector Database [Database]** | **Stores text embeddings for semantic search** |
| **Software** | **Backend[API]** | **Serves as the main interface between the system's front-end and back-end components** |
| **Software** | **Gmail API [API]** | **Fetches documents sent through emails** |
| **Software** | **Mobile Application [Android]** | **Provides user access and interaction through an Android interface** |
| **Software** | **Web Application [Website]** | **Provides a user interface for accessing the system via the web** |
| **Hardware** | **Scanner** | **Used by the company to scan documents and send them into the system** |

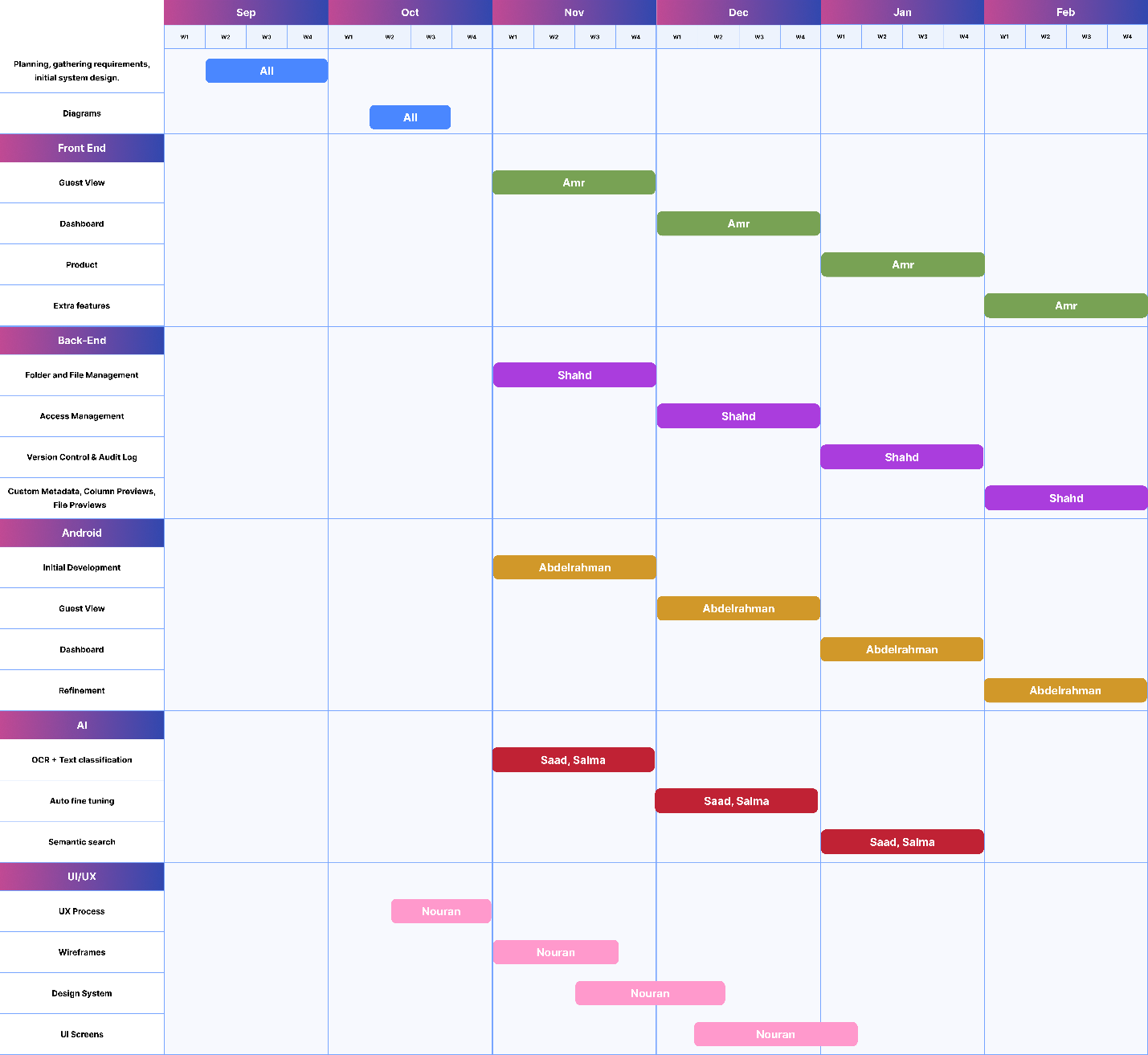
1. **Question 6: Compared to the existing solutions that you have read about, what are the new ideas that you intend to develop and include in your solution?**  
   Our product differentiates itself through several innovative features that leverage AI for enhanced document management:  
   1)AI Models for Document Detection and Classification  
   2)Semantic Search Capabilities  
   3)Comprehensive Version Control  
   4)Granular Access Management  
   5)AI-Driven Optical Character Recognition (OCR)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Feature | Describe of feature | DOKMEE | DocuWare | M-Files | ArchivAI |
| Folder Creation and Initial Setup | The system lets users manually organize files, creating a structure for AI to learn from. | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Don’t Have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill |
| Automatic File Categorization with AI | After setup, the AI automatically sorts new files based on content, eliminating manual organization. | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |
| Semantic Search for Easy Document Retrieval | We will implement semantic search to allow users to find documents by meaning, enhancing search speed and accuracy. | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill |
| Text Extraction with OCR | The system will use OCR to extract text from scanned documents, enabling searches for non-digital content. | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |
| File Parsing and Metadata for Better Search | Each document will be scanned for key details, saved as metadata, and used to simplify searches. | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |
| Testing and Feedback | We will test the AI for accurate categorization and smooth searches, using user feedback to improve the system. | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Don’t have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill |
| Deployment and Ongoing Updates | After testing, the system will be deployed for users, with continuous updates based on feedback and technological advancements. | Have this feature  Badge Cross with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill | Have this feature  Checkmark with solid fill |

1. **Question 7: Describe your teams system development methodology and quality assurance process.**

**Methodology  
  
Quality Assurance**



1. **Question 8: Give your initial project management plans, especially: distribution of responsibilities and tasks, milestone schedule, training plan, risk management, contingency plans, and change control procedures.**  
   **Milestone Schedule**  
   ****

**Training Plan**A screenshot of a computer

Description automatically generated

**Contringency Plan**

**A table with text on it

Description automatically generated**

**Distribution of responsibilites and tasks**A screenshot of a spreadsheet

Description automatically generated

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk** | **Likelihood** | **Impact** | **Risk Response** |
| **System Downtime** | **Medium** | **High** | **Implement redundancy and failover mechanisms; have a dedicated technical team for quick recovery.** |
| **Data Loss** | **Medium** | **High** | **Schedule regular backups; store data securely with encryption; have a data recovery plan.** |
| **Security Breach** | **Medium** | **High** | **Implement strong access controls, and regular security audits** |
| **AI Model Downtime or Malfunction** | **Medium** | **Medium** | **Automatically switching to a previous checkpoint model; Monitor model performance continuously** |
| **Classification Uncertainty** | **High** | **Medium** | **Flag Low-confidence predictions for review.** |
| **Data Mismatch or Inconsistency** | **Medium** | **Medium** | **Establish data validation rules; use automated checks to detect inconsistencies.** |
| **Communication Failure** | **Low** | **Low** | **Set up alternate communication channels; have a clear internal and external communication plan.** |

**Change Control Procedure**

