i. Machine Learning Project Scoping:

1. Major Components in the System (3 marks):
- Data Collection Module: This component is responsible for capturing 3D scans of individuals' faces, ensuring high-quality and standardized data acquisition.
- Preprocessing and Feature Extraction: This step involves cleaning, normalizing, and extracting relevant features from the 3D scans to prepare them for machine learning.
- Machine Learning Model: The core of the system, where a machine learning algorithm is trained and deployed to classify whether a person has an infectious disease or not.
- User Interface: A user-friendly interface for scientists or medical professionals to interact with the system, input data, and view results.

2. Key Components Requiring Machine Learning (3 marks):
- Feature Extraction: Machine learning methods will be used to extract meaningful features from the 3D scans, such as facial landmarks, texture patterns, and other relevant information.
- Classification Model: Supervised learning methods like Convolutional Neural Networks (CNNs) or Support Vector Machines (SVMs) can be employed for disease classification based on the extracted features.
- Model Optimization: Techniques like hyperparameter tuning and model selection will be used to optimize the performance of the classification model.

3. Data Collection for the Training Set (2 marks):
- Data can be collected by partnering with medical institutions and hospitals to acquire 3D scans of individuals diagnosed with infectious diseases.
- Ensure diverse and representative samples across various demographics, ages, and disease types.
- Augment the dataset if necessary to balance classes and increase model robustness.

4. Testing Machine Learning Components (2 marks):
- Cross-Validation: Perform k-fold cross-validation to assess the model's generalization and prevent overfitting.
- Holdout Testing: Reserve a portion of the dataset for independent testing to evaluate the model's performance on unseen data.
- Example: Divide the dataset into 80% training and 20% testing. Train the model on the training data, and then evaluate its accuracy, precision, recall, and F1-score on the testing data.

5. Extending the System (2 marks):

- Real-time Monitoring: Implement real-time scanning and decision-making for immediate disease detection.
- Continuous Learning: Incorporate mechanisms for the model to continuously adapt and improve as new data becomes available.
- Integration with Healthcare Systems: Connect the system with healthcare databases to enhance disease tracking and reporting capabilities.

ii. Ethics in AI and Data Science:

1. Privacy and Ethical Issues (4 marks):

- Informed Consent: Ensure individuals are informed about the use of their facial data for disease detection and obtain their consent.
- Data Security: Implement robust data encryption and access controls to prevent unauthorized use or data breaches.
- Bias and Fairness: Regularly audit and mitigate biases in the system to avoid discrimination against certain demographic groups.
- Transparency: Make the system's decision-making process transparent, and provide explanations for classification outcomes to build trust.

2. Advantages and Disadvantages (4 marks):

Advantages:

- Improved Disease Detection: Enhances public health efforts by identifying infectious individuals quickly.
- Public Safety: Helps contain disease outbreaks by monitoring potential carriers in public spaces.
- Automation: Reduces the burden on human resources for manual disease screening.

Disadvantages:

- Privacy Concerns: Raises issues regarding the surveillance and tracking of individuals without their explicit consent.
- False Positives/Negatives: ML models may make errors, leading to unnecessary panic or missed cases.
- Ethical Dilemmas: Balancing public health benefits with individual privacy rights and ethical considerations.
- Dependence on Technology: Over-reliance on AI may undermine traditional healthcare practices and human judgment.

Please note that while this response outlines a framework, the actual implementation and ethical considerations should be thoroughly reviewed, and privacy and ethical experts should be consulted for specific

guidance.