**Table of Contents**

1. **Introduction**  
   1.1 Overview  
   1.2 Problem Statement  
   1.3 Scope and Objectives  
   1.4 Report Organization  
   1.5 Work Methodology  
   1.6 Work Plan (Gantt Chart)
2. **Related Work (Literature Review)**  
   2.1 Speech Emotion Recognition (SER)  
   2.2 LLM‑based Candidate Scoring and Retrieval  
   2.3 Existing Online Interview Platforms  
   2.4 Summary of Findings

**Part I – Speech Emotion Recognition (SER)**

1. **Datasets and Preprocessing**  
   3.1 Sources (RAVDESS, CREMA‑D, TESS, SAVEE)  
   3.2 Data Cleaning and Labeling  
   3.3 Feature Extraction (MFCC, ZCR, RMS)  
   3.4 Data Augmentation (noise, stretching, shifting, pitch)
2. **Model Architecture and Training**  
   4.1 CNN Architecture Overview  
   4.2 Training Pipeline (hyperparameters, callbacks)  
   4.3 Validation and Testing Procedures  
   4.4 Performance Metrics and Confusion Matrix
3. **Emotion Analyzer Module**  
   5.1 Audio Extraction from Video (FFmpeg)  
   5.2 Segmentation Strategy and Feature Pipeline  
   5.3 Model Loading and Prediction  
   5.4 Result Aggregation (dominant emotion, confidence)  
   5.5 Integration with Streamlit Interface

**Part II – Answer Evaluation**

1. **Rubric Generation and Scoring Criteria**  
   6.1 Technical and HR Rubrics (Azure OpenAI)  
   6.2 Example Criteria (Clarity, Accuracy, Professionalism)  
   6.3 Storage Format (JSON)
2. **Dataset Collection for Q&A**  
   7.1 HR and Technical Question Banks  
   7.2 Historical Answers and Scores  
   7.3 CSV/JSON Structure
3. **Retrieval and Matching Mechanism**  
   8.1 Embeddings via Azure/OpenAI  
   8.2 FAISS Vector Indexes  
   8.3 Exact‑Match Retrieval  
   8.4 Relevance Search with LangChain
4. **Evaluation Pipeline**  
   9.1 Handling “I don’t know” and Short Answers  
   9.2 Rubric-Based Scoring via GPT‑4o  
   9.3 Combining Historical Scores with New Evaluations  
   9.4 Batch Processing of Multiple Q&A Pairs  
   9.5 Output Format (JSON Results)

**System Integration**

1. **Video Interview Application**  
   10.1 Project Structure Overview (Streamlit app, components, utils)  
   10.2 Audio/Video Recording Workflow  
   10.3 Transcription with Whisper  
   10.4 Grammar Checking (LanguageTool + GPT)  
   10.5 Result Aggregation and PDF Report Generation
2. **Trials and Experiments**  
   11.1 SER Model Training Experiments  
   11.2 Rubric Generation Trials  
   11.3 Alternative Whisper Models  
   11.4 Embedding and Vector Database Tests
3. **Evaluation and Results**  
   12.1 SER Accuracy and Confusion Matrix  
   12.2 Example Answer Evaluations  
   12.3 Sample Interview Sessions  
   12.4 Human Evaluation (if applicable)
4. **Analysis and Discussion**  
   13.1 Emotion Distribution Findings  
   13.2 Candidate Scoring Trends and Insights
5. **Conclusions and Future Work**  
   14.1 Summary of Achievements  
   14.2 Potential Improvements and Extensions
6. **References**
7. **Appendices**
   * Additional Code Snippets or Training Logs
   * Extra Figures and Tables
   * Configuration Files and Paths