

## AI Pitch Analysis Model Report

### Objective

The objective of this assignment is to develop an LLM-powered pitch analysis model that evaluates startup pitch decks by providing a pitch score, highlighting strengths, and identifying weaknesses.

### Dataset & Inputs

- A set of sample pitch decks (sample\_pitch\_deck\_1.pdf, sample\_pitch\_deck\_2.pdf, sample\_pitch\_deck\_3.pdf).
- Key sections to be analyzed: Problem, Solution, Market, Business Model, Financials, Team.

### Tasks

#### 1. Text Extraction & Preprocessing

- **Method:** PDF parsing using PyPDF2 to extract text content from each pitch deck.
- **Preprocessing:** Removing unnecessary elements and formatting the extracted text for analysis.

#### 2. Feature Engineering

- **Section Identification:** Extracting key sections (Problem, Solution, Market, Business Model, Financials, Team) from the text.
- **Weight Assignment:** Assigning weights to sections based on their importance for evaluation.

#### 3. Scoring Model

- **Model Used:** Fine-tuned BERT model for sequence classification.
- **Process:**
  - Tokenize text using BertTokenizer.
  - Generate predictions using BertForSequenceClassification.
  - Convert model output to scores using Softmax and scale the results between 0 and 100.
- **Metrics:** Scoring is based on the relevance, clarity, and completeness of each section.

#### 4. Strength & Weakness Analysis

- **Feedback Generation:**
  - If a section score  $\geq 80$ : "Excellent - This section is well-crafted and provides relevant information."
  - If a section score  $\geq 60$ : "Good - This section is decent but could use some improvements."
  - If a section score  $< 60$ : "Needs Improvement - This section lacks clarity or detail."

## 5. Output

- **Visual Representation:** Bar chart of section-wise scores.
- **Results File:** pitch\_analysis\_results.json containing scores and feedback.

### Analysis Results (Sample Pitch Decks)

#### Sample Pitch Deck 1:

- Problem: 59 (Needs Improvement)
- Solution: 59 (Needs Improvement)
- Market: 57 (Needs Improvement)
- Business Model: 55 (Needs Improvement)
- Financials: 56 (Needs Improvement)
- Team: 54 (Needs Improvement)

#### Sample Pitch Deck 2:

- Problem: 60 (Needs Improvement)
- Solution: 58 (Needs Improvement)
- Market: 58 (Needs Improvement)
- Business Model: 56 (Needs Improvement)
- Financials: 57 (Needs Improvement)
- Team: 56 (Needs Improvement)

#### Sample Pitch Deck 3:

- Problem: 63 (Needs Improvement)
- Solution: 60 (Needs Improvement)
- Market: 58 (Needs Improvement)
- Business Model: 58 (Needs Improvement)
- Financials: 57 (Needs Improvement)
- Team: 56 (Needs Improvement)

### Insights & Recommendations

- The BERT model is effective in identifying and scoring various sections of pitch decks.
- Most sections require improvement, indicating a need for enhancing clarity, detail, and presentation quality.
- Future work could involve using more advanced models like GPT/Gemini API for even better analysis.

### Output:--

Analyzing sample\_pitch\_deck\_1.pdf...

Section Scores:

Problem: 0.59

Solution: 0.59

Market: 0.57

Business Model: 0.55

Financials: 0.56

Team: 0.54

### Feedback:

Problem: Needs Improvement - This section lacks clarity or detail.

Solution: Needs Improvement - This section lacks clarity or detail.

Market: Needs Improvement - This section lacks clarity or detail.

Business Model: Needs Improvement - This section lacks clarity or detail.

Financials: Needs Improvement - This section lacks clarity or detail.

Team: Needs Improvement - This section lacks clarity or detail.

