AI-Powered Lead Scoring – Technical & Business Report

Business Problem

Sales teams waste time chasing leads that don't convert. Our tool uses AI to **rank leads by quality**, helping teams focus on **high-potential contacts** based on **job title seniority**, **company size**, and **email validity**. This streamlines prospecting and improves outreach efficiency.

Model Summary

- **Architecture**: 3-layer MLP (Multi-Layer Perceptron) in PyTorch
- **Input Features**:
    - GloVe 50D average embedding of job title
    - Normalized company size (0–1)
    - Binary indicator for valid email
- **Output**: A score between 0–1 representing lead quality
- **Training Set**: 200 synthetically generated leads with heuristic labels (e.g., C-suite or company size > 100 = good lead)

Data Processing

1. **Job Title Embedding**: Mean of GloVe vectors for each word
2. **Company Size**: Normalized to 0–1 scale
3. **Email Validity**: Regex-based validation → 1 or 0
4. **Deduplication**: Leads with duplicate emails are removed before scoring

Performance

- Scores align closely with business intuition:
    - CEOs, VPs at large companies → ~1.00
    - Developers at small companies → <0.6
- Model remains small, fast, and suitable for **real-time lead prioritization**
- No external API needed; runs entirely offline

Deployment & UX

- Simple **Streamlit web app**:
    - Upload CSV of leads
    - View deduplicated, AI-scored leads
    - Download results for CRM upload