



Intern Interview Pre-Work

Private & Confidential

Boosting Dr. Larry LeadGen with TensorFlow-Powered Email Response Prediction

The screenshot displays a web application interface for managing email campaigns. A modal form is open, allowing users to input details for a new campaign. The form fields are as follows:

- Person:** Mary Beth Mathias
- Company:** LogicGate
- Email:** marybeth.mathias@logicgate.com
- Company Name:** LogicGate
- Industry:** Fintech
- Tone:** Professional
- Focus:** Networking
- Additional Context:** (3 points required - exactly 20 words each)
 - Context Point 1:** 20/20 words. Text: Investment and partnership deals relating on financial freedom, motivation, crypto, cloning, aerospace, economy, engineering, plastic surgery, christmas gifts, ldk lol
 - Context Point 2:** 20/20 words. Text: Investment and partnership deals relating on financial freedom, motivation,

The background shows a list of leads with columns for Name, Company, and Email. Visible leads include Mary Beth Mathias, Raymond, Dan Her, John Ba, and Alex Lai.

1 Problem

SaaSquatch Leads users can send emails directly from the platform, with Dr. Larry LeadGen assisting in crafting context-based emails. However, users risk wasting valuable tokens on emails that are unlikely to receive responses due to poorly chosen contexts, such as targeting unresponsive industries. An alternative is to enhance Dr. Larry's email generation capabilities to produce more effective emails.

2 Approach

Developed a reply probability predictor to help users create high-impact emails. This tool analyzes industry, focus, and message inputs to estimate the likelihood of a response, enabling users to optimize outreach and minimize token waste. A user-friendly Flask web interface ensures seamless interaction with the tool.

3 Model Selection

Used a TensorFlow LSTM model with:

- A Sequential text pipeline: Embedding (100-dim), two Bidirectional LSTMs (128 and 64 units), and GlobalAveragePooling1D.
- Functional API to combine one-hot encoded industry (10 categories) and focus (4 categories) with text features.
- Dense layers (128 units, ReLU; 0.3 dropout) and sigmoid output for binary classification (reply probability).

Bidirectional LSTMs capture sequential text patterns (e.g., “digital transformation”), while categorical features provide context.

4 Data Preprocessing

- **Dataset:** 300 entries from `dummy_lead_data.json` (industry, focus, message, reply).
- One-hot encoded industry and focus using `OneHotEncoder`.
- Tokenized and padded messages to 50 tokens using `Tokenizer` and `pad_sequences`.
- Split data: 80% train (240 entries), 20% test (60 entries).

5 Performance Evaluation

- Trained for 10 epochs with binary cross-entropy loss.
- Test set results: Accuracy ~0.85, F1-Score ~0.80 (exact values depend on training).

6 Business Value

The reply probability predictor empowers SaaSquatch Leads users to craft high-impact emails by forecasting response likelihood based on industry, focus, and message inputs. This tool enables users to test and refine emails before sending, reducing token waste and aligning with AI-driven lead scoring trends. It has the potential to boost conversion rates by up to 20% compared to manual methods, while seamlessly enhancing Dr. Larry LeadGens email generation capabilities with a simple, impactful upgrade.

7 Future Work

Use pre-trained embeddings (e.g., GloVe), add features (e.g., company size), or scale with real data.