AUGUST 24, 2025

Improving Customer Support with Text Representations

Case Study

SUJIT DHANUKA

Background

In 2024, **HelpNow**, a mid-sized SaaS company, faced challenges in scaling its **customer support operations**. The company's support tickets, submitted through email and chat, were piling up—leading to delays in resolving customer issues.

The VP of Operations wanted to use **machine learning** to:

1. Automatically **categorize support tickets** into predefined categories (Billing, Technical Issue, Account Access, General Query).
2. Identify **duplicate tickets** to reduce agent workload.
3. Provide **semantic search** so agents could find relevant past solutions quickly.

The company engaged a team of data science interns to test different **text representation methods** and evaluate their performance.

The Tasks

Task 1: Ticket Categorization

- **Objective**: Assign each incoming ticket to one of the four categories.
- **Methods**:
  - One-Hot Encoding + Logistic Regression
  - TF-IDF + Naïve Bayes
  - Word Embeddings (pretrained GloVe/Word2Vec) + Neural Network
- **Performance Metric**: Classification accuracy, Precision/Recall per category, F1-score.

Task 2: Duplicate Ticket Detection

- **Objective**: Identify whether a new ticket is a duplicate of an already submitted one.
- **Methods**:
  - TF-IDF vectors + Cosine Similarity
  - SVD (Latent Semantic Analysis) + Cosine Similarity
  - Word Embeddings (average pooling) + Cosine Similarity
- **Performance Metric**: ROC-AUC, Average Precision, Recall@K (how often the correct duplicate is in the top K).

Task 3: Semantic Search for Resolutions

- **Objective**: Given a new ticket, retrieve the most relevant historical ticket resolutions.
- **Methods**:
  - TF-IDF Index
  - SVD (dimensionality reduction of TF-IDF)
  - Word Embedding Similarity Search
- **Performance Metric**: Mean Reciprocal Rank (MRR), Normalized Discounted Cumulative Gain (NDCG).

Task 4: Pretrained Models (Advanced Exploration)

- **Objective**: Test modern transfer learning models on ticket classification and semantic search.
- **Methods**:
  - **ULMFiT (Universal Language Model Fine-Tuning)**: Fine-tuned on support tickets.
  - **DistilBERT**: Lightweight transformer for classification and semantic similarity.
- **Performance Metric**: Comparison with traditional methods (accuracy, latency, resource usage).

Task 5: Hierarchical Neural Networks

- **Objective**: Move beyond flat representations to structured encodings.

- **Approach**:
    - **Character → Word → Sentence Encoding** using a Hierarchical NN.
    - Extendable to **multimodal encoding** (Word + Image + Audio) for future chatbot design.
- **Performance Metric**: Improvement in F1/Recall on noisy text (typos, slang), interpretability of higher-level encodings.

Text Representation & Model Approaches

1. **One-Hot Encoding** → Baseline, very sparse.
2. **TF-IDF** → Simple, interpretable, strong baseline.
3. **SVD / LSA** → Captures latent semantic structure.
4. **Word Embeddings (Word2Vec/GloVe)** → Dense semantic vectors.
5. **ULMFiT** → Transfer learning with language model fine-tuning.
6. **DistilBERT** → Transformer-based pretrained model, efficient and scalable.
7. **Hierarchical Neural Networks** → Captures context at multiple levels (char → word → sentence).

Decision Point

By the end of the pilot, **HelpNow's leadership team** must decide:

- Should they deploy **classical methods (TF-IDF/SVD)** that are cost-efficient and interpretable?
- Or should they invest in **pretrained models (DistilBERT/ULMFiT)** for higher accuracy but more compute?
- Should they explore **hierarchical/multimodal encodings** as a long-term R&D direction?

As interns, you will analyze the performance trade-offs and **recommend the best approach** balancing accuracy, interpretability, and scalability.