#ASLI ENGINEERING

Natural Language Search at Zomato

SWIPE

BY

ARPIT BHAYANI

Zomato is a restaurant aggregator that allows

- restaurants to list themselves
- people to order from their favourite restaurant

One of the main medium to discover food/restaurant is Search!

On zomato people search for 3 things

1. Dish ⟶ pizza, burger
2. Restaurant ⟶ Dominos, Pizza Hut
3. Cuisine ⟶ Chinese

" Pizza Dominos "
↑

People also fire complex search queries that combine the above entities.

Simple Search Engines : Tokenize, Match, Rank

Typical configuration: Title match >>> Description

Consider the query: " Best coffee near me "

This will bring up restaurants having "Best" in their name

- Best Coffee Cafe
- Best Bliss
- Bar Best

Hence, we cannot do lexical match, instead we need

" Natural Language Understanding "

ARPIT BHAYANI

Plus because of voice search the queries are more verbose

- Garlic Bread with Cheese Dip

- Veg restaurant in koramangola

- Chai and Samosa

Single-intent queries are easy to answer, but such queries require
deeper understanding of domain & natural language.

Understanding Intent

Given a query, understand what user is seeking

Dominos outlets near me

Restaurant ‿ Location

Intent: Nearest Outlet
Entity: Dominos

Zomato classifies search queries in 3 categories

1. Dish + Dish : chai and samosa

2. Restaurant + Dish : mcd burger

3. Restaurant/Dish + near me, best, irrelevant text : pizza near me

Best Dominos Pizza
near me
↑

It becomes really easy to understand the query
once we have narrowed the scope.

ARPIT BHAYANI

Challenges in query understanding

1. Unavailability of labelled data

2. queries involving more than one language

   - Makni Dal ke saath naan
   - Sabse acha pizza

   } Multi-lingual queries

3. phrases that mean the same

   - Mcd burger
   - Mcd k burger

   } Conjunctions (language agnostic) and fillers

4. spelling variations

   - Rajma rice   Rajma chawal
   - Rumali roti   Roomali Roti

   } Synonym and Phonetics

Model training

```
┌─────────────────┐
│  ┌───────────┐  │
│  │ word2vec  │  │
│  └───────────┘  │        ┌──────────┐        ┌───────┐
│                 │ ───→   │ Bi-LSTM  │ ───→   │  CRF  │
│  ┌───────────┐  │        └──────────┘        └───────┘
│  │Byte Pair Enc│ │
│  └───────────┘  │
└─────────────────┘
```
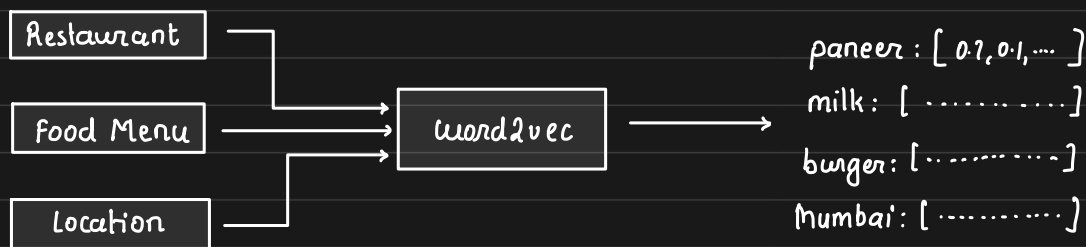
word2vec → Neural network model to learn word associations

1. we can provide a corpus and train the model

2. it generates vector for each term in vocabulary

These vectors can then be used with other algorithms

Zomato trained word2vec on their domain data

```
┌──────────────┐
│  Restaurant  │ ─────┐
└──────────────┘      │          ┌──────────────┐
┌──────────────┐      ├────────► │   word2vec   │ ──────►
│  Food Menu   │ ─────┤          └──────────────┘
└──────────────┘      │
┌──────────────┐      │
│  Location    │ ─────┘
└──────────────┘
```

paneer : [0.7, 0.1, ...]

milk : [··········]

burger : [··········]

Mumbai : [··········]

Byte-pair encoding for tokenization

Instead of tokenizing text generically, BPE is a supervised tokenization.

It is a sub-word tokenizer
⇓
it extracts token from within the
word as well.

FRIEDRICE ⟶ FRIED , RICE

Because these words are prevelent
in the corpus.

This helps in extracting the max information
out from the text and query.

ARPIT BHAYANI

Bidirectional LSTM + CRF to do sequence taggin

These tokens are used as vocab to generate word embeddings
and then a Neural Network is trained to understand Named Entities

Mix veg sabji aur roti → | Mix veg sabji | aur | roti |

                                            Dish   Dish

Jack's Aaloo Tikki Burger → | Jack's | Aaloo Tikki Burger |

                                 Restaurant   Dish

Architecture

Depending on the intent,
  Search Query to be fired gets adjusted

Search   ④   Search Engine

②

Gateway

Restaurant

Food Menu

Location

③

Model

Extracting info
from query to understand
the intent and
then fire Search

ARPIT BHAYANI