Named Entity Disambiguation Boosted With Knowledge Graph

Brian Lin
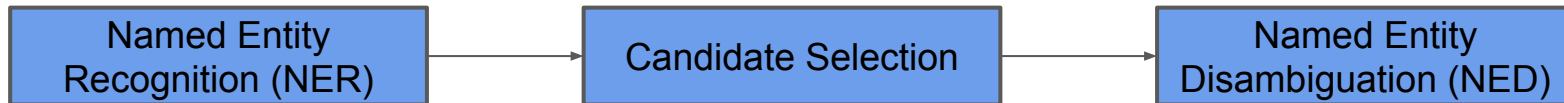Shane Ong
Cory Williams
Yun Bin (Matteo) Zhang

Overview
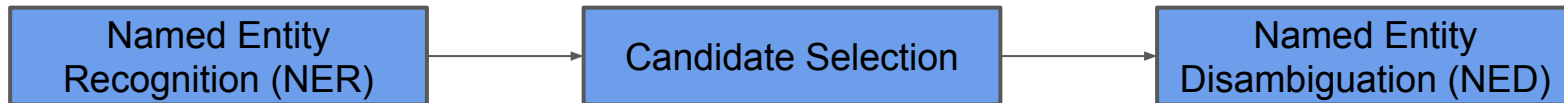
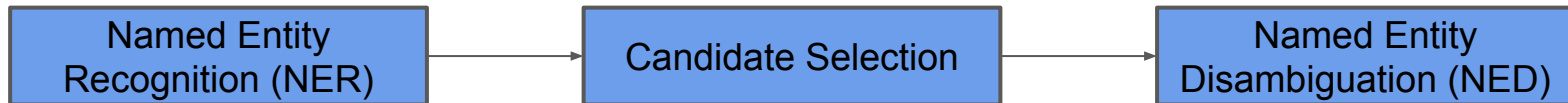| Named Entity Recognition (NER)	→	Candidate Selection	→	Named Entity Disambiguation (NED)

Knowledge Base:
- Jaguar Cars (car)
- Jaguar Land Rover (company)
- Jaguar (animal)

I want to drive my Jaguar around the city.

Overview

```
┌────────────────────┐     ┌────────────────────┐     ┌────────────────────┐
│   Named Entity     │ ──→ │ Candidate Selection│ ──→ │   Named Entity     │
│  Recognition (NER) │     │                    │     │ Disambiguation (NED)│
└────────────────────┘     └────────────────────┘     └────────────────────┘
```

Knowledge Base:
- Jaguar Cars (car)
- Jaguar Land Rover (company)
- Jaguar (animal)

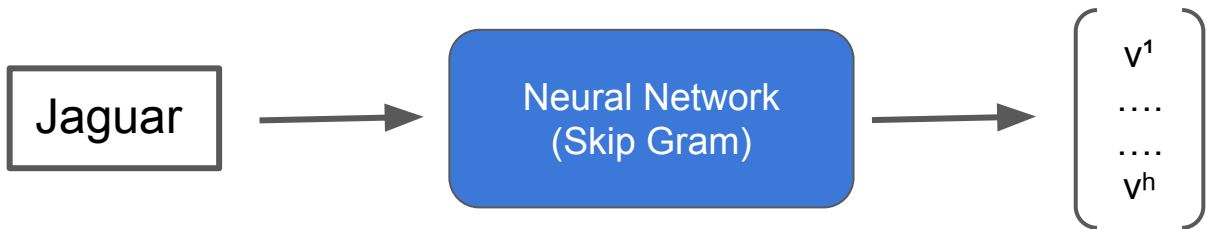I want to drive my Jaguar around the city.

Overview

Named Entity Recognition (NER) → Candidate Selection → Named Entity Disambiguation (NED)

Knowledge Graph:

Toyota Prius (car) —Competitor of→ Jaguar Cars (car) —Produced by→ Jaguar Land Rover (company)

Jaguar Land Rover (company) —Vehicle Assembly plant in→ Brazil (country)

Jaguar Land Rover (company) —Subsidiary of→ Tata Motors (company)

Jaguar (animal) —Native to→ Brazil (country)
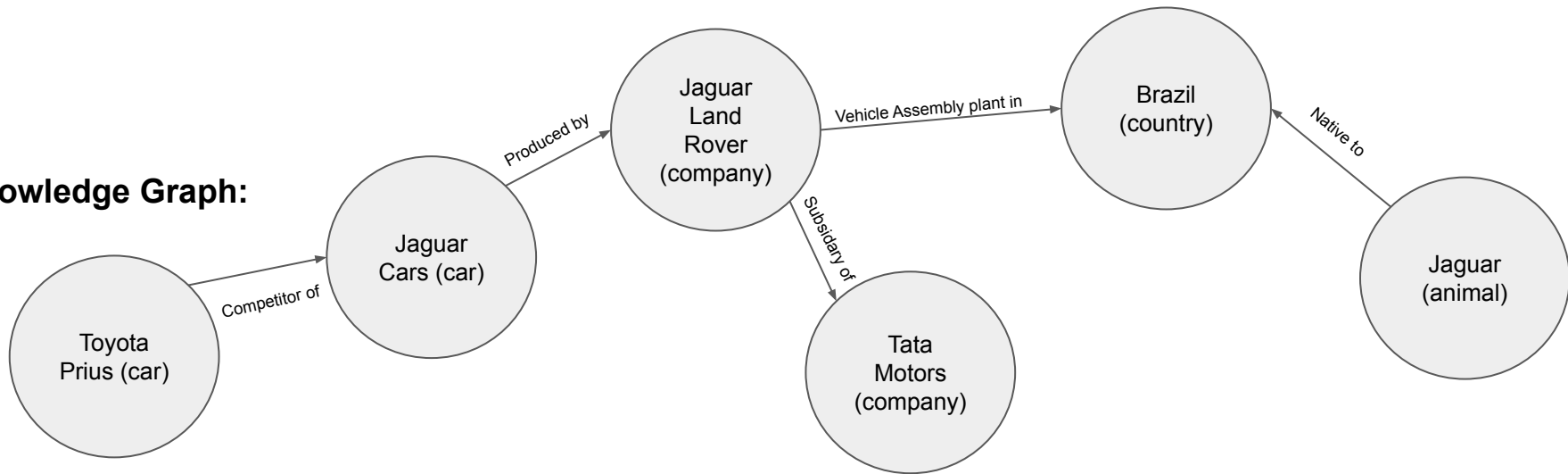
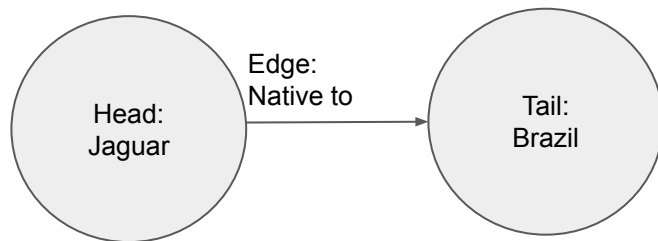… I want to drive my Jaguar around the city. … It is so much better than my previous Toyota Prius

Word Embeddings

- Vector representations to capture context of a word in a document
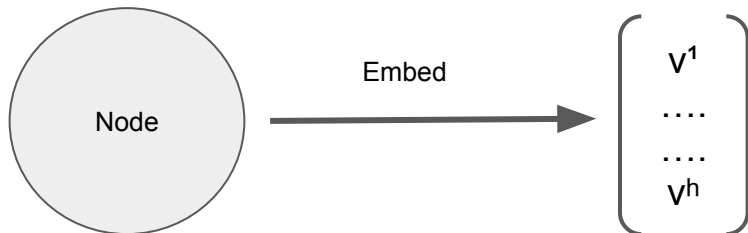
- Words with similar context should be spatially close

```
┌──────────┐        ┌──────────────────┐        ⎡ v¹  ⎤
│  Jaguar  │ ─────> │  Neural Network  │ ─────> │ ....│
└──────────┘        │   (Skip Gram)    │        │ ....│
                    └──────────────────┘        ⎣ vʰ  ⎦
```

$$\begin{bmatrix} v^1 \\ .... \\ .... \\ v^h \end{bmatrix}$$

Knowledge Graph Embeddings

- Millions of triplets stored as (head, edge, tail)

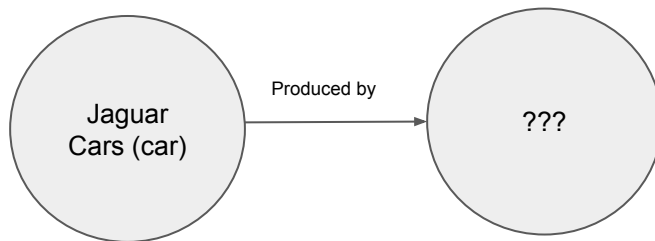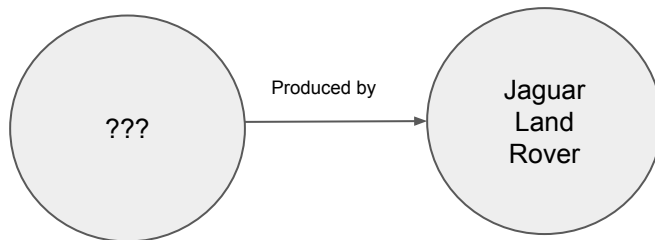- Interested in the underlying structural information:

Knowledge Graph Embeddings

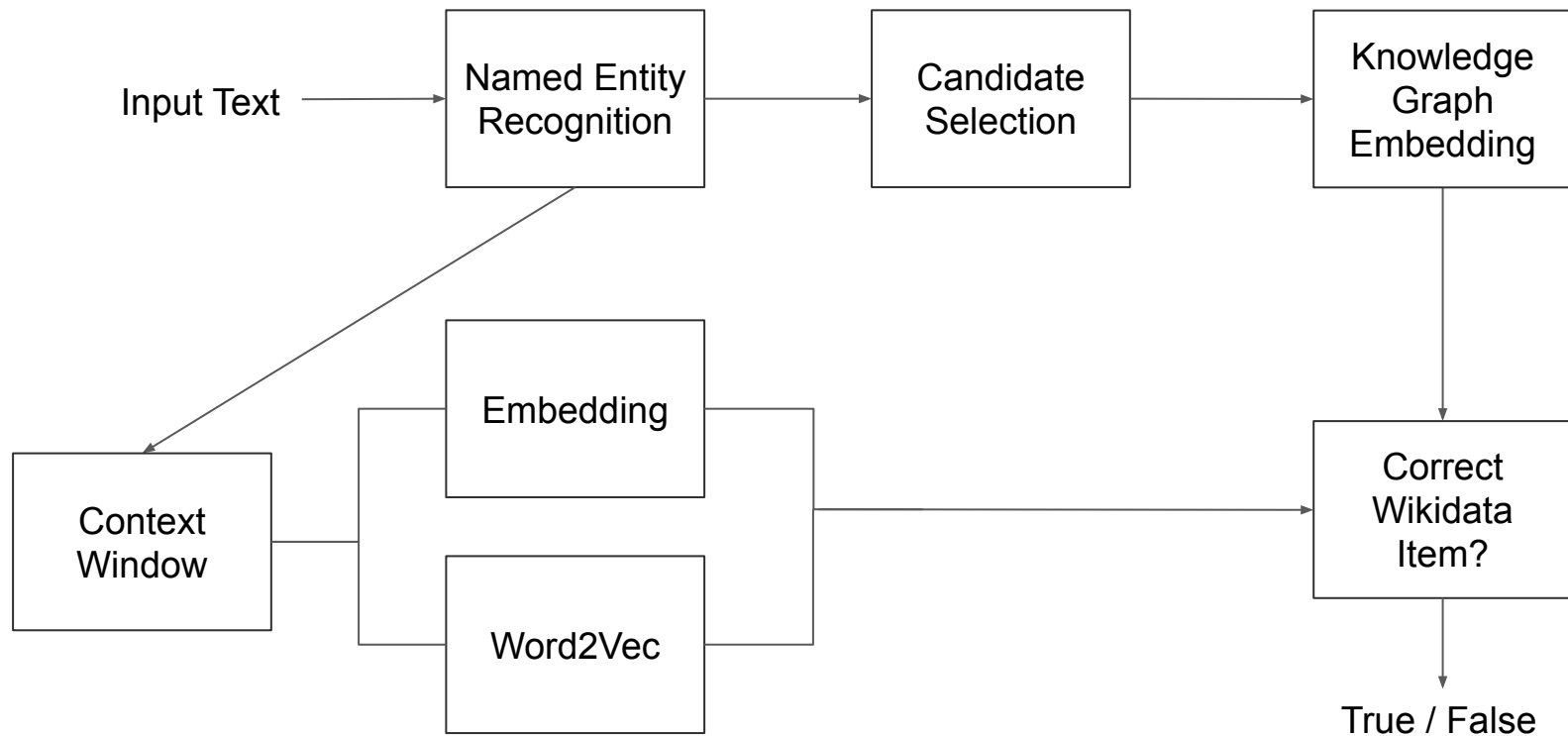Trained on the task of predicting:

- Tail in (head, edge, ???)

- Head in (???, edge, tail)

Model 1

Model 1
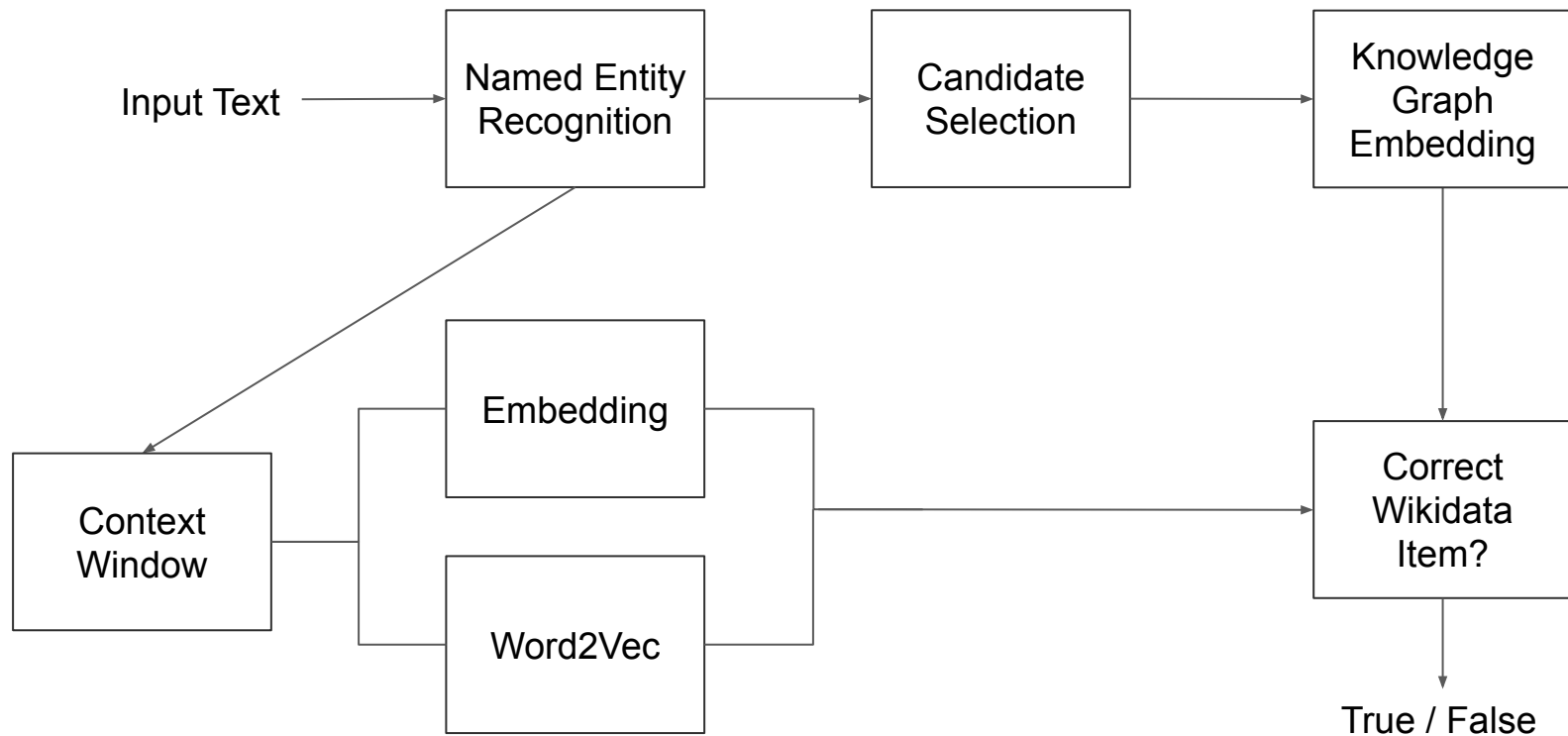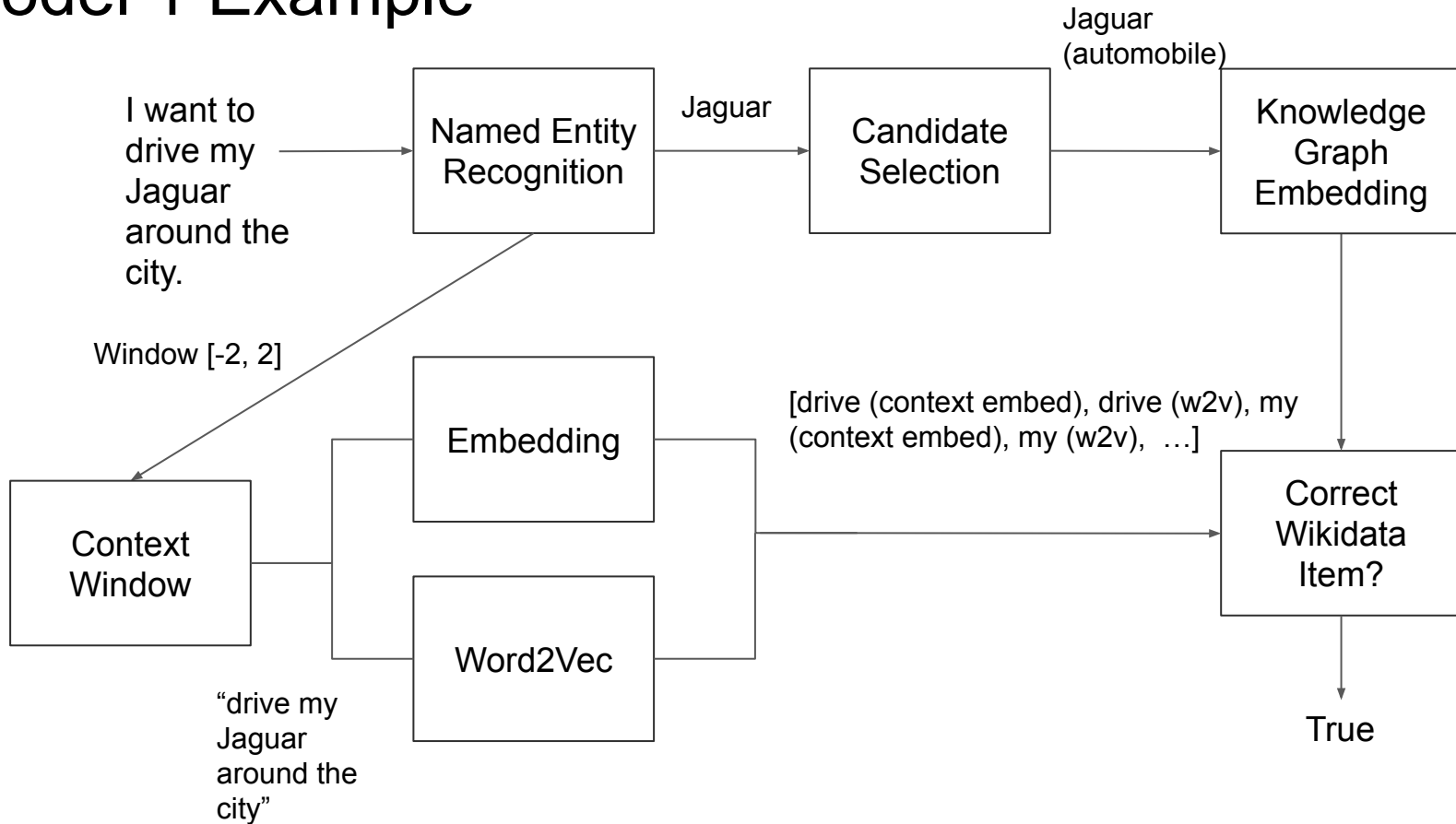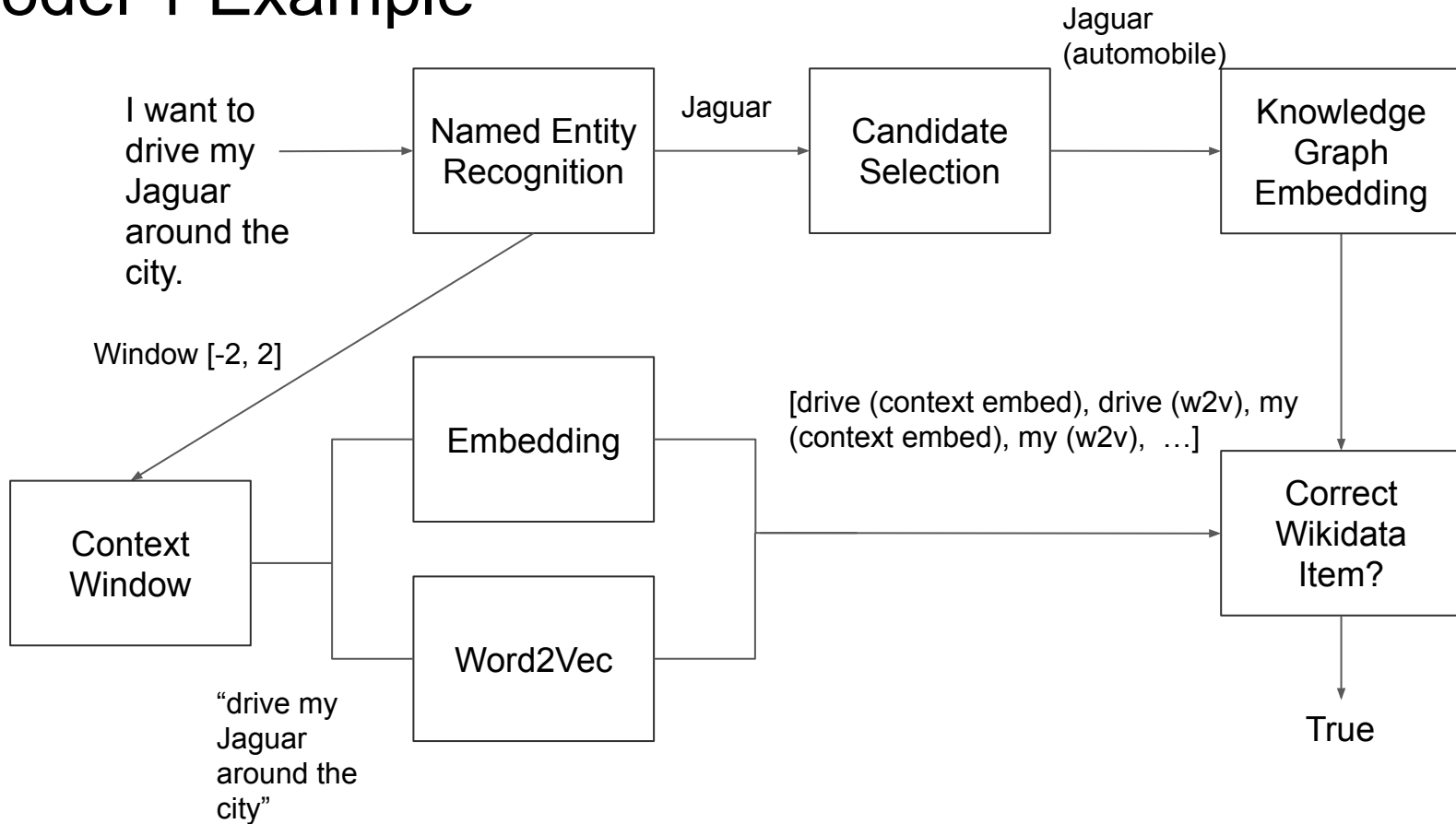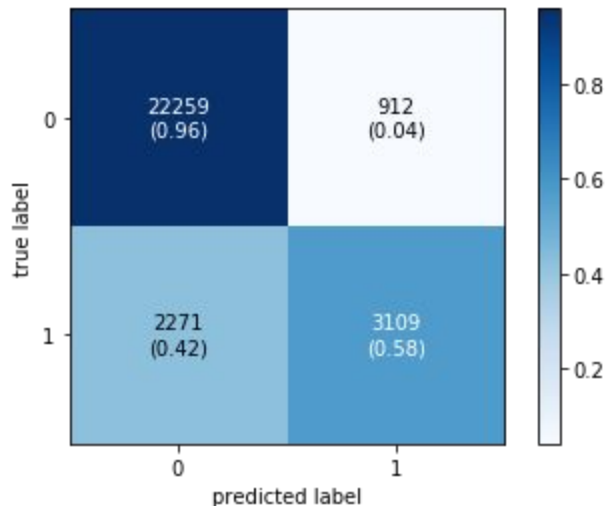
Model 1 Example

I want to drive my Jaguar around the city.

Named Entity Recognition

Jaguar

Candidate Selection

Jaguar (automobile)

Knowledge Graph Embedding

Window [-2, 2]

Context Window

"drive my Jaguar around the city"

Embedding

Word2Vec

[drive (context embed), drive (w2v), my (context embed), my (w2v), …]

Correct Wikidata Item?

True

Model 1 Example

I want to drive my Jaguar around the city.

Named Entity Recognition

Jaguar

Candidate Selection

Jaguar (automobile)

Knowledge Graph Embedding

Window [-2, 2]

Context Window

Embedding

Word2Vec

[drive (context embed), drive (w2v), my (context embed), my (w2v), …]

Correct Wikidata Item?

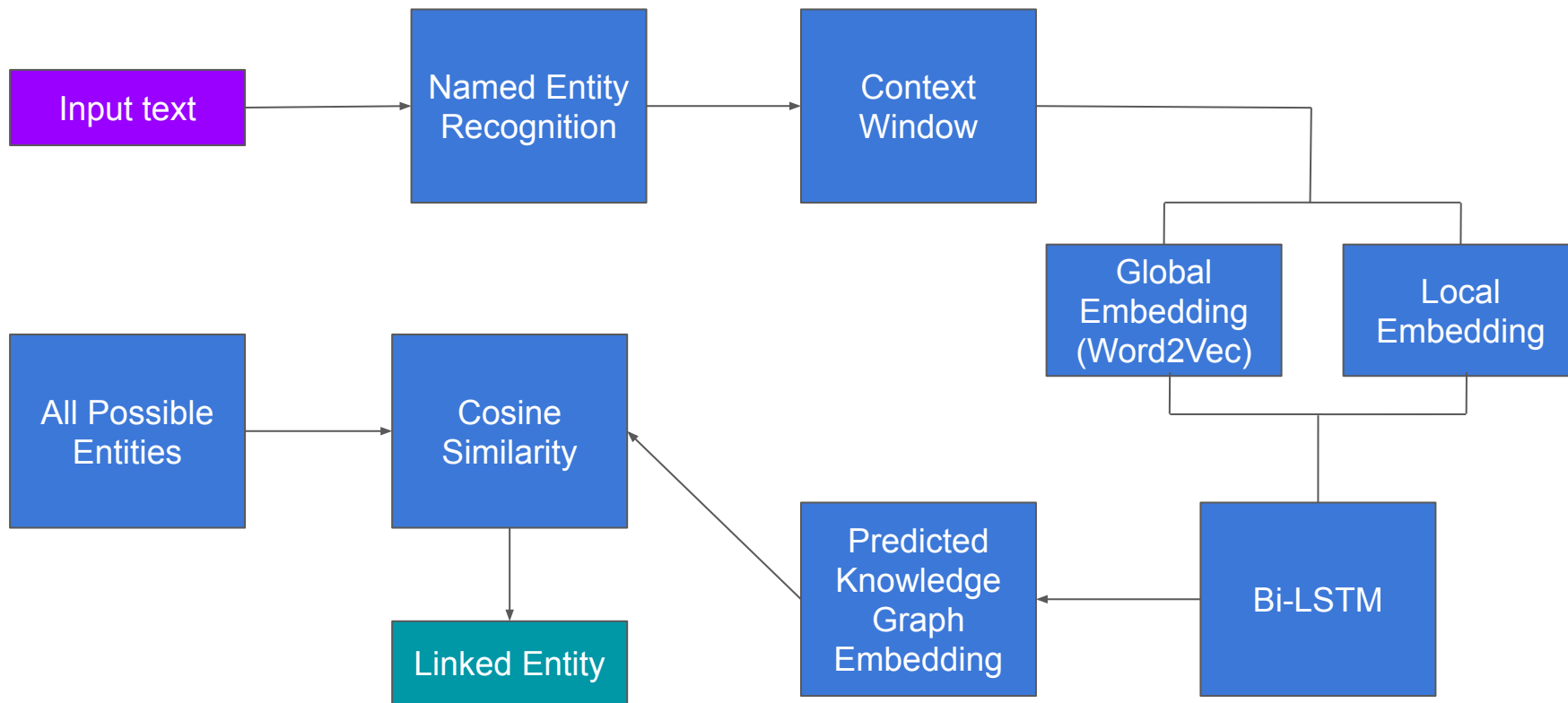"drive my Jaguar around the city"

True

Simplified Model 1 Results

- Sample Dataset: 114204 rows
  - True entity with positive label, false entity with negative label
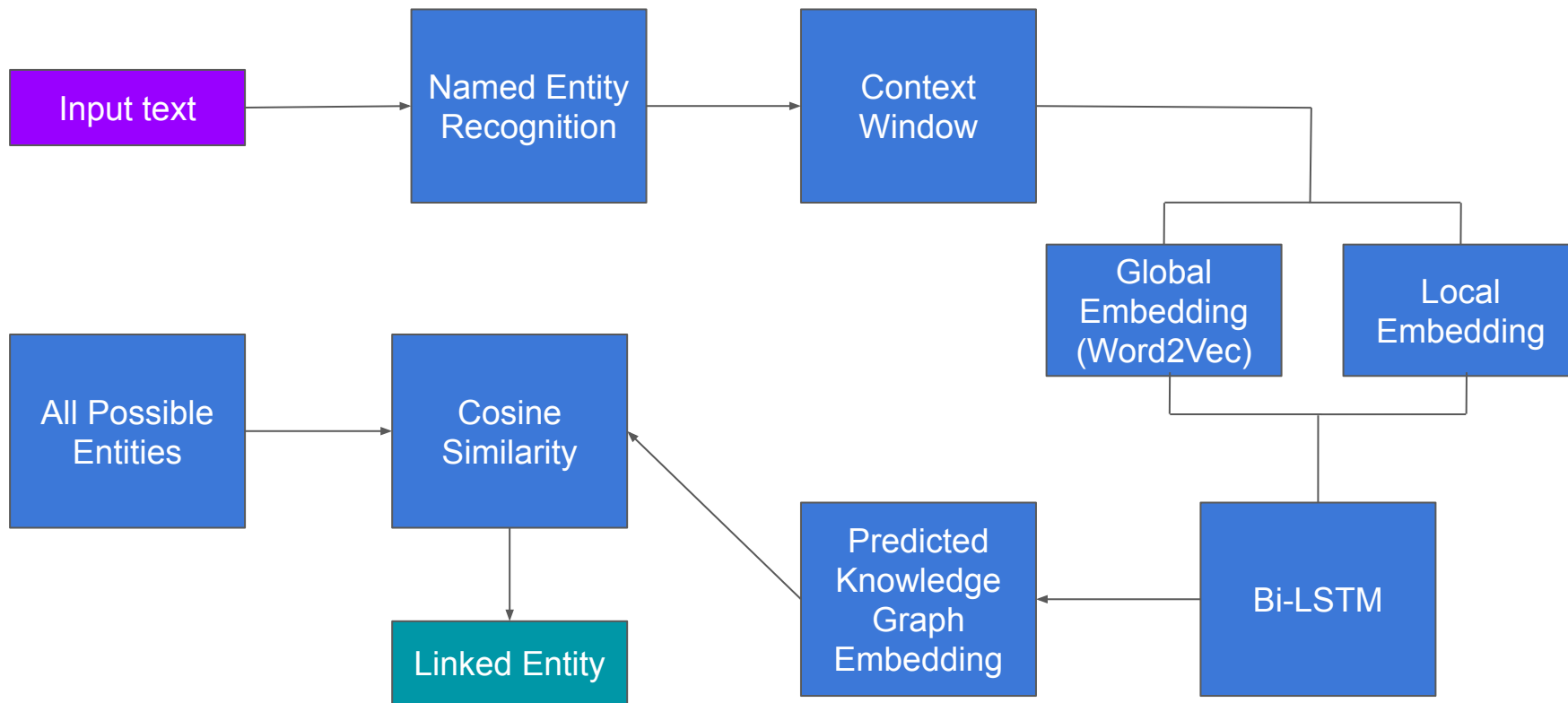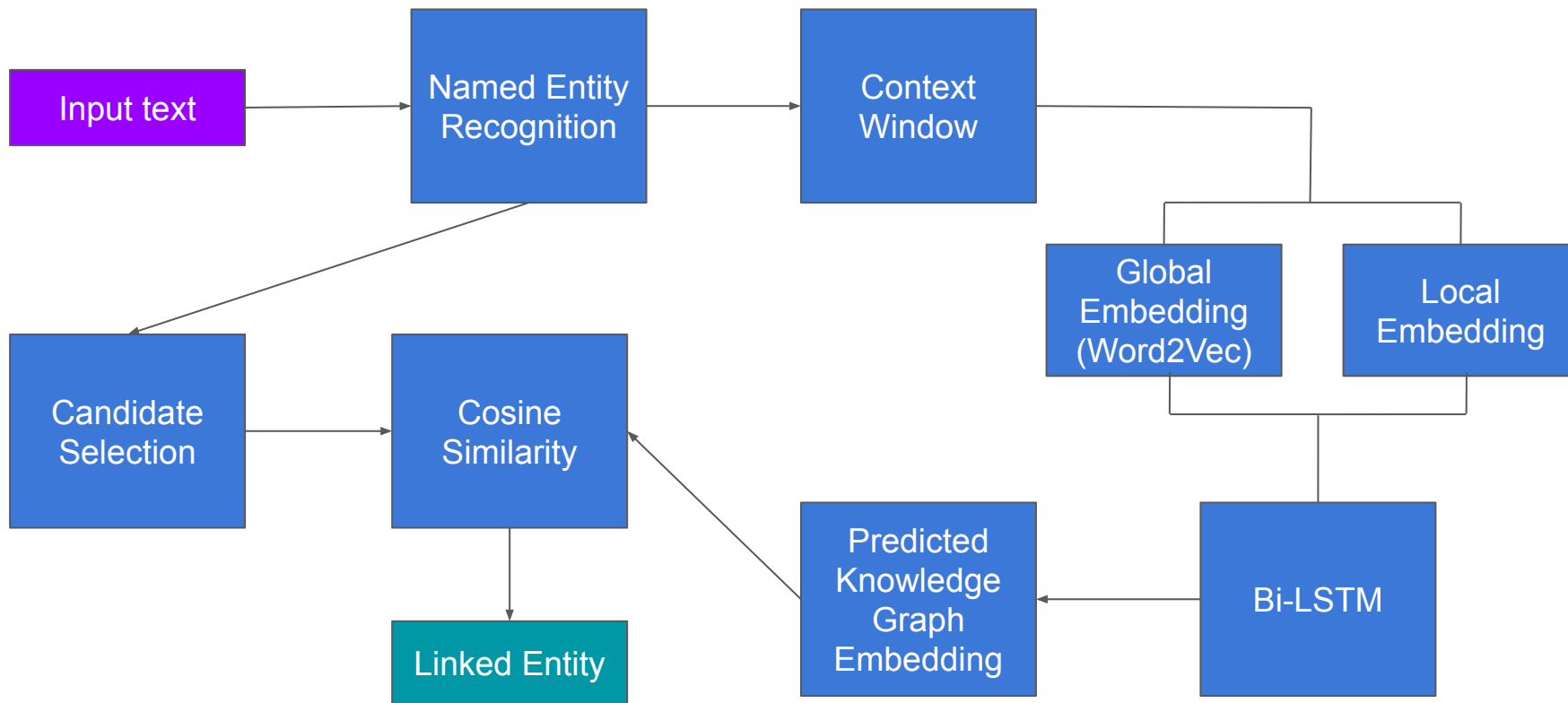 - 18.84% positive, 81.16% negative labels
  - Test accuracy: 88.85% (2 Hidden Layer NN)

Model 2

Model 2

Model 2

Model 2