# Natural Language Processing
## for Question-Answer Systems

Jack Arnold | Prakrit Duangsutha | Caleb Berger | Jatin Jain | Rashmi Ananth

## What is NLP?

### Basic Concept

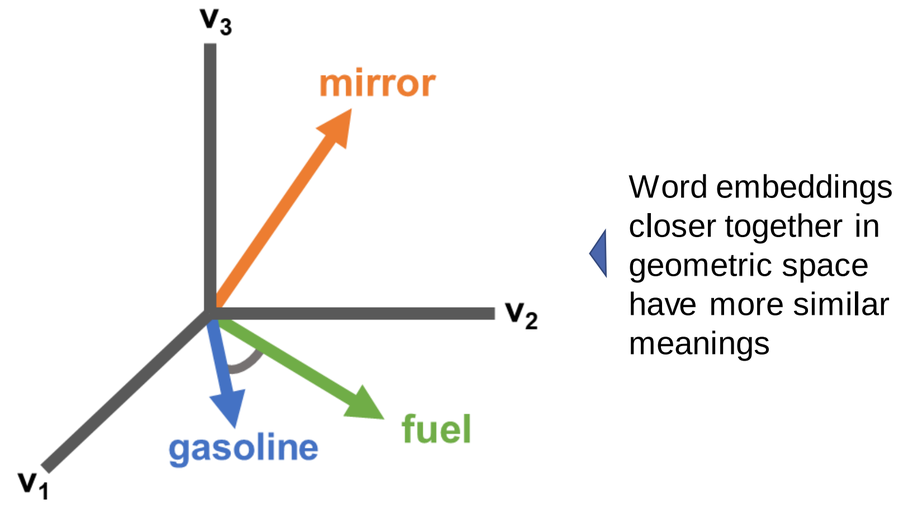
- Train a computer to interpret human language
- Represent words as numbers (high dimensional vectors) known as "embeddings"
- Vectors close to each other are interpreted as similar

### Common Applications

- Languages translation
- Question Answer (QA) Systems
- Text Prediction

### Challenges

- Languages are very ambiguous
- Sentences have more complex meanings




## Attention and Transformers

### Mimicking Human Approach

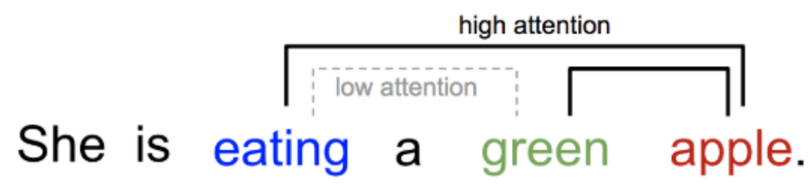
- When we (humans) interpret sentences, we focus on the most important words
- Similarly, a model using attention only uses the parts of input where most relevant information is concentrated
- Weights are computed using a neural network

### "Attention is All You Need"

- Self-attention is used so a model knows how words relate to one another
- It can keep track of long-term word dependences among sentences

### Breakthrough Success

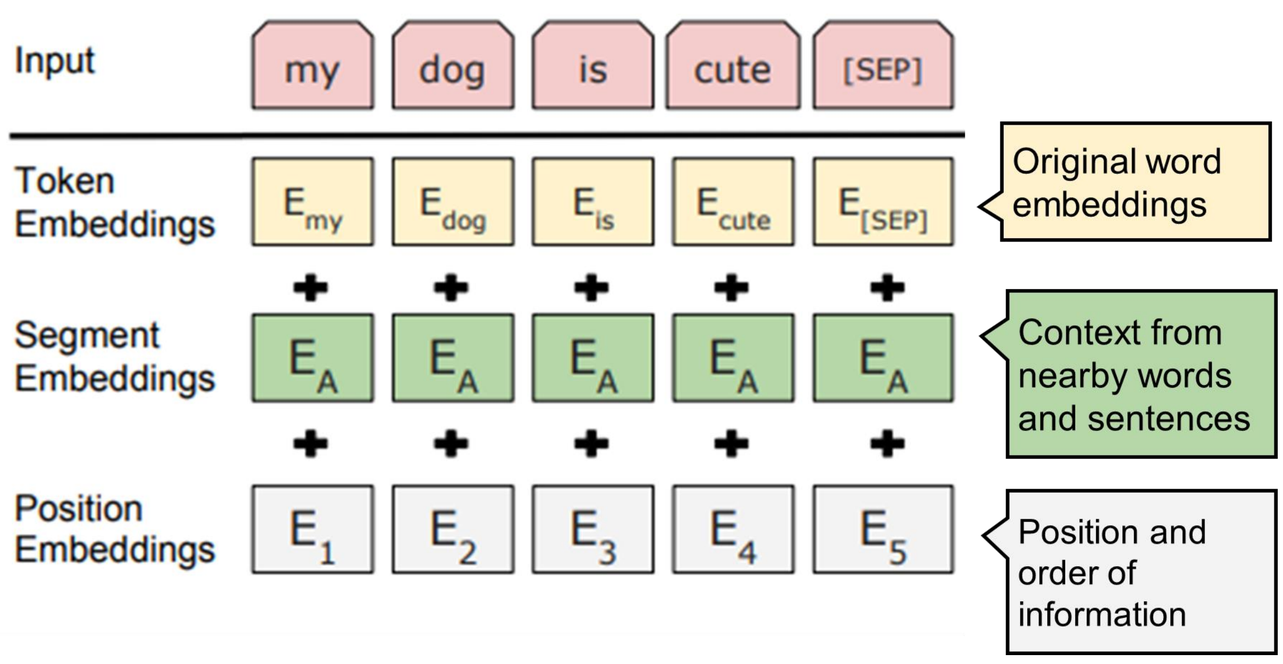
- Attention produces much better results with record setting accuracy
- Almost all state-of-the-art models use attention
- Model designs like Transformers utilize attention for accuracy and provide higher speed by being entirely feed-forward



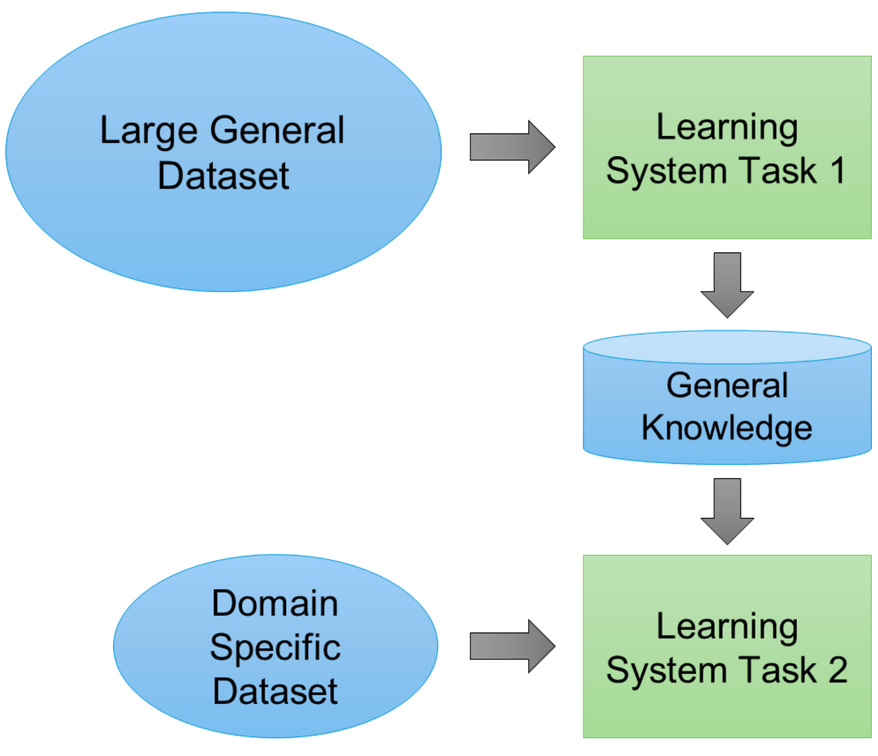

## Modern NLP Architecture

### BERT
### (Bidirectional Encoder Representations from Transformers)

- Consider both the meaning of words (in context) and their significance
- Rather than reading left to right, the model considers all surrounding words to embed context (bidirectional)
- A small change to existing models to greatly improve accuracy for many tasks
- Important especially for QA systems to locate relevant answers




## Transfer Learning




### Repurposing a Model

- Training an NLP model for every dataset is computationally expensive
- Transfer learning uses models already trained on large datasets to significantly reduce computation
- We can start by training our QA system based on pre-trained word vectors
- We then use domain specific data (related to Ford vehicles) to fine tune the model for the task at hand