

## NLP Assignment 4

### 1. Mention the steps that go into resolving any NLP issue.

- Understand the problem and its context.
- Gather and preprocess the data relevant to the problem.
- Choose appropriate NLP techniques and models.
- Train and fine-tune the models based on the data.
- Evaluate the models and iterate for improvements.

### 2. Mention how important word2vec is.

Word2Vec is crucial because it transforms words into vectors, preserving semantic relationships and context. This facilitates meaningful mathematical operations on words, aiding in tasks like similarity measurement, sentiment analysis, and more, advancing the accuracy and efficiency of NLP models.

### 3. Explain the process of feature extraction in NLP.

Feature extraction in NLP involves converting raw text into numerical or structured features that machine learning models can use. Techniques like Bag of Words, TF-IDF, and Word Embeddings (e.g., Word2Vec) transform text into features suitable for model training.

### 4. What is the distinction between precision and recall?

**Precision** measures the accuracy of positive predictions among all predicted positives. It focuses on minimizing false positives.

**Recall** measures the accuracy of positive predictions among all actual positives. It focuses on minimizing false negatives.

### 5. Explain the concept of tokenization.

Tokenization is the process of breaking down text into smaller units, or tokens, such as words, phrases, symbols, or other meaningful elements. It's a fundamental step in NLP that facilitates further analysis, processing, and understanding of the text.

**6. Mention the distinction between formal and informal language.**

**Formal Language:**

Typically used in professional, academic, or official settings. It follows specific rules of grammar, structure, and vocabulary, and aims for precision and clarity.

**Informal Language:**

Used in casual, everyday conversations. It's less structured, often includes slang, colloquialisms, and is adapted to the context and relationship between speakers. Informal language may prioritize expression over strict grammatical rules.