

Q1

<p>"As the evening settled in, the world seemed to slow down. The golden hue of the setting sun spilled across the rooftops, painting everything in a warm glow. Leaves rustled softly in the breeze, and the faint sound of a flute echoed from a nearby balcony, blending into the quiet hum of life winding down. Children's laughter faded into the distance, replaced by the occasional bark of a dog or the clang of utensils from kitchens preparing dinner. There was a sense of calm, of a day well-lived and a night waiting with promises of dreams, stories, and stillness."</p>	<p>'As', 'the', 'evening', 'settled', 'in', ',', 'the', 'world', 'seemed', 'to', 'slow', 'down', '.', 'The', 'golden', 'hue', 'of', 'the', 'setting', 'sun', 'spilled', 'across', 'the', 'rooftops', ',', 'painting', 'everything', 'in', 'a', 'warm', 'glow', '.', 'Leaves', 'rustled', 'softly', 'in', 'the', 'breeze', ',', 'and', 'the', 'faint', 'sound', 'of', 'a', 'flute', 'echoed', 'from', 'a', 'nearby', 'balcony', ',', 'blending', 'into', 'the', 'quiet', 'hum', 'of', 'life', 'winding', 'down', '.', 'Children', 's', 'laughter', 'faded', 'into', 'the', 'distance', ',', 'replaced', 'by', 'the', 'occasional', 'bark', 'of', 'a', 'dog', 'or', 'the', 'clang', 'of', 'utensils', 'from', 'kitchens', 'preparing', 'dinner', '.', 'There', 'was', 'a', 'sense', 'of', 'calm', ',', 'of', 'a', 'day', 'well', '-', 'lived', 'and', 'a', 'night', 'waiting', 'with', 'promises', 'of', 'dreams', ',', 'stories', ',', 'and', 'stillness', '.'</p>	<p>Tokenization</p>
<p>"As the evening settled in, the world seemed to slow down. The golden hue of the setting sun spilled across the rooftops, painting everything in a warm glow. Leaves rustled softly in the breeze, and the faint sound of a flute echoed from a nearby balcony, blending into the quiet hum of life winding down. Children's laughter faded into the distance, replaced by the occasional bark of a dog or the clang of utensils from kitchens preparing dinner. There was a sense of calm, of a day well-lived and a night waiting with promises of dreams, stories, and stillness."</p>	<p>[evening, settled, ,, world, slow, ,, golden, hue, setting, sun, spilled, rooftops, ,, painting, warm, glow, ,, Leaves, rustled, softly, breeze, ,, faint, sound, flute, echoed, nearby, balcony, ,, blending, quiet, hum, life, winding, ,, Children, laughter, faded, distance, ,, replaced, occasional, bark, dog, clang, utensils, kitchens, preparing, dinner, ,, sense, calm, ,, day,-, lived, night, waiting, promises, dreams, ,, stories, ,, stillness, .</p>	<p>Stop word removal</p>

<p>"As the evening settled in, the world seemed to slow down. The golden hue of the setting sun spilled across the rooftops, painting everything in a warm glow. Leaves rustled softly in the breeze, and the faint sound of a flute echoed from a nearby balcony, blending into the quiet hum of life winding down. Children's laughter faded into the distance, replaced by the occasional bark of a dog or the clang of utensils from kitchens preparing dinner. There was a sense of calm, of a day well-lived and a night waiting with promises of dreams, stories, and stillness."</p>	<p>['As', ',', 'as', ',', 'As']</p> <p>['the', ',', 'the', ',', 'the']</p> <p>['evening', ',', 'even', ',', 'evening']</p> <p>['settled', ',', 'settl', ',', 'settled']</p> <p>['in', ',', 'in', ',', 'in']</p> <p>[',', ',', ', , ', ', , ', ', , ']</p> <p>['the', ',', 'the', ',', 'the']</p> <p>['world', ',', 'world', ',', 'world']</p> <p>['seemed', ',', 'seem', ',', 'seemed']</p> <p>['to', ',', 'to', ',', 'to']</p> <p>['slow', ',', 'slow', ',', 'slow']</p> <p>['down', ',', 'down', ',', 'down']</p> <p>[',', ',', ', , ', ', , ', ', , ']</p> <p>['The', ',', 'the', ',', 'The']</p> <p>['golden', ',', 'golden', ',', 'golden']</p> <p>['hue', ',', 'hue', ',', 'hue']</p> <p>['of', ',', 'of', ',', 'of']</p> <p>['the', ',', 'the', ',', 'the']</p> <p>['setting', ',', 'set', ',', 'setting']</p> <p>['sun', ',', 'sun', ',', 'sun']</p> <p>['spilled', ',', 'spill', ',', 'spilled']</p> <p>['across', ',', 'across', ',', 'across']</p> <p>['the', ',', 'the', ',', 'the']</p> <p>['rooftops', ',', 'rooftop', ',', 'rooftop']</p> <p>[',', ',', ', , ', ', , ', ', , ']</p> <p>['painting', ',', 'paint', ',', 'painting']</p> <p>['everything', ',', 'everyth', ',', 'everything']</p> <p>['in', ',', 'in', ',', 'in']</p> <p>['a', ',', 'a', ',', 'a']</p> <p>['warm', ',', 'warm', ',', 'warm']</p> <p>['glow', ',', 'glow', ',', 'glow']</p> <p>[',', ',', ', , ', ', , ', ', , ']</p> <p>['Leaves', ',', 'leav', ',', 'Leaves']</p> <p>['rustled', ',', 'rustl', ',', 'rustled']</p> <p>['softly', ',', 'softli', ',', 'softly']</p> <p>['in', ',', 'in', ',', 'in']</p> <p>['the', ',', 'the', ',', 'the']</p> <p>['breeze', ',', 'breez', ',', 'breeze']</p>	<h2>Stemming vs Lemmatization</h2>
--	---	------------------------------------

	<p>[',, ', ', ', ', ', ']</p> <p>['and', ', ', 'and', ', ', 'and']</p> <p>['the', ', ', 'the', ', ', 'the']</p> <p>['faint', ', ', 'faint', ', ', 'faint']</p> <p>['sound', ', ', 'sound', ', ', 'sound']</p> <p>['of', ', ', 'of', ', ', 'of']</p> <p>['a', ', ', 'a', ', ', 'a']</p> <p>['flute', ', ', 'flute', ', ', 'flute']</p> <p>['echoed', ', ', 'echo', ', ', 'echoed']</p> <p>['from', ', ', 'from', ', ', 'from']</p> <p>['a', ', ', 'a', ', ', 'a']</p> <p>['nearby', ', ', 'nearbi', ', ', 'nearby']</p> <p>['balcony', ', ', 'balconi', ', ', 'balcony']</p> <p>[',, ', ', ', ', ', ']</p> <p>['blending', ', ', 'blend', ', ', 'blending']</p> <p>['into', ', ', 'into', ', ', 'into']</p> <p>['the', ', ', 'the', ', ', 'the']</p> <p>['quiet', ', ', 'quiet', ', ', 'quiet']</p> <p>['hum', ', ', 'hum', ', ', 'hum']</p> <p>['of', ', ', 'of', ', ', 'of']</p> <p>['life', ', ', 'life', ', ', 'life']</p> <p>['winding', ', ', 'wind', ', ', 'winding']</p> <p>['down', ', ', 'down', ', ', 'down']</p> <p>[',, ', ', ', ', ', ']</p> <p>['Children', ', ', 'children', ', ', 'Children']</p> <p>['s', ', ', 's', ', ', 's']</p> <p>['laughter', ', ', 'laughter', ', ', 'laughter']</p> <p>['faded', ', ', 'fade', ', ', 'faded']</p> <p>['into', ', ', 'into', ', ', 'into']</p> <p>['the', ', ', 'the', ', ', 'the']</p> <p>['distance', ', ', 'distanc', ', ', 'distance']</p> <p>[',, ', ', ', ', ', ']</p> <p>['replaced', ', ', 'replac', ', ', 'replaced']</p> <p>['by', ', ', 'by', ', ', 'by']</p> <p>['the', ', ', 'the', ', ', 'the']</p> <p>['occasional', ', ', 'occasion', ', ', 'occasional']</p> <p>['bark', ', ', 'bark', ', ', 'bark']</p> <p>['of', ', ', 'of', ', ', 'of']</p> <p>['a', ', ', 'a', ', ', 'a']</p>	
--	--	--

	['dog', 'ˌ', 'dog', 'ˌ', 'dog'] ['or', 'ˌ', 'or', 'ˌ', 'or'] ['the', 'ˌ', 'the', 'ˌ', 'the'] ['clang', 'ˌ', 'clang', 'ˌ', 'clang'] ['of', 'ˌ', 'of', 'ˌ', 'of'] ['utensils', 'ˌ', 'utensil', 'ˌ', 'utensil'] ['from', 'ˌ', 'from', 'ˌ', 'from'] ['kitchens', 'ˌ', 'kitchen', 'ˌ', 'kitchen'] ['preparing', 'ˌ', 'prepar', 'ˌ', 'preparing'] ['dinner', 'ˌ', 'dinner', 'ˌ', 'dinner'] [ˌ, ˌ, ˌ, ˌ, ˌ] ['There', 'ˌ', 'there', 'ˌ', 'There'] ['was', 'ˌ', 'wa', 'ˌ', 'wa'] ['a', 'ˌ', 'a', 'ˌ', 'a'] ['sense', 'ˌ', 'sens', 'ˌ', 'sense'] ['of', 'ˌ', 'of', 'ˌ', 'of'] ['calm', 'ˌ', 'calm', 'ˌ', 'calm'] [ˌ, ˌ, ˌ, ˌ, ˌ] ['of', 'ˌ', 'of', 'ˌ', 'of'] ['a', 'ˌ', 'a', 'ˌ', 'a'] ['day', 'ˌ', 'day', 'ˌ', 'day'] ['well', 'ˌ', 'well', 'ˌ', 'well'] [ˌ, ˌ, ˌ, ˌ, ˌ] ['lived', 'ˌ', 'live', 'ˌ', 'lived'] ['and', 'ˌ', 'and', 'ˌ', 'and'] ['a', 'ˌ', 'a', 'ˌ', 'a'] ['night', 'ˌ', 'night', 'ˌ', 'night'] ['waiting', 'ˌ', 'wait', 'ˌ', 'waiting'] ['with', 'ˌ', 'with', 'ˌ', 'with'] ['promises', 'ˌ', 'promis', 'ˌ', 'promise'] ['of', 'ˌ', 'of', 'ˌ', 'of'] ['dreams', 'ˌ', 'dream', 'ˌ', 'dream'] [ˌ, ˌ, ˌ, ˌ, ˌ] ['stories', 'ˌ', 'stori', 'ˌ', 'story'] [ˌ, ˌ, ˌ, ˌ, ˌ] ['and', 'ˌ', 'and', 'ˌ', 'and'] ['stillness', 'ˌ', 'still', 'ˌ', 'stillness'] [ˌ, ˌ, ˌ, ˌ, ˌ]	
--	---	--

<p>"As the evening settled in, the world seemed to slow down. The golden hue of the setting sun spilled across the rooftops, painting everything in a warm glow. Leaves rustled softly in the breeze, and the faint sound of a flute echoed from a nearby balcony, blending into the quiet hum of life winding down. Children's laughter faded into the distance, replaced by the occasional bark of a dog or the clang of utensils from kitchens preparing dinner. There was a sense of calm, of a day well-lived and a night waiting with promises of dreams, stories, and stillness."</p>	<p>As CONJ the DET evening NOUN settled VERB in ADP , PUNCT the DET world NOUN seemed VERB to PART slow VERB down ADP . PUNCT The DET golden ADJ hue NOUN of ADP the DET setting VERB sun NOUN spilled VERB across ADP the DET rooftops NOUN , PUNCT painting VERB everything PRON in ADP a DET warm ADJ glow NOUN . PUNCT</p>	<p>POS tagging</p>
--	--	--------------------

	Leaves NOUN rustled VERB softly ADV in ADP the DET breeze NOUN , PUNCT and CCONJ the DET faint ADJ sound NOUN of ADP a DET flute NOUN echoed VERB from ADP a DET nearby ADJ balcony NOUN , PUNCT blending VERB into ADP the DET quiet ADJ hum NOUN of ADP life NOUN winding VERB down ADP . PUNCT Children NOUN 's PART	
--	--	--

	<p>laughter NOUN</p> <p>faded VERB</p> <p>into ADP</p> <p>the DET</p> <p>distance NOUN</p> <p>, PUNCT</p> <p>replaced VERB</p> <p>by ADP</p> <p>the DET</p> <p>occasional ADJ</p> <p>bark NOUN</p> <p>of ADP</p> <p>a DET</p> <p>dog NOUN</p> <p>or CCONJ</p> <p>the DET</p> <p>clang NOUN</p> <p>of ADP</p> <p>utensils NOUN</p> <p>from ADP</p> <p>kitchens NOUN</p> <p>preparing VERB</p> <p>dinner NOUN</p> <p>. PUNCT</p> <p>There PRON</p> <p>was VERB</p> <p>a DET</p> <p>sense NOUN</p> <p>of ADP</p> <p>calm NOUN</p> <p>, PUNCT</p> <p>of ADP</p>	
--	---	--

	a DET day NOUN well ADV - PUNCT lived VERB and CCONJ a DET night NOUN waiting VERB with ADP promises NOUN of ADP dreams NOUN , PUNCT stories NOUN , PUNCT and CCONJ stillness NOUN . PUNCT	
--	--	--

Q2

Which Captures Semantics Better and Why?

1. Bag of Words (BoW)

- Captures: Word frequency only.
- Limitations:
 - Ignores context and word order.
 - Treats words like "good" and "great" as unrelated.
- Conclusion: Very limited in capturing semantics.

2. TF-IDF (Term Frequency – Inverse Document Frequency)

- Captures: Importance of a word in a document relative to the whole corpus.

- Strengths:
 - Reduces the weight of common words.
 - More informative than BoW in filtering noise.
- Limitations:
 - Still ignores meaning and context.
 - “happy” and “joyful” are still treated as different.
- Conclusion: Better than BoW in weighting, but still poor at capturing semantics.

3. Word Embeddings (Word2Vec, GloVe, BERT, etc.)

- Captures: Contextual meaning and semantic relationships between words.
- Strengths:
 - Words with similar meanings have similar vector representations.
 - e.g., $\text{vec}(\text{"king"}) - \text{vec}(\text{"man"}) + \text{vec}(\text{"woman"}) \approx \text{vec}(\text{"queen"})$
 - Can capture analogies, context, and relationships.
 - Contextual embeddings (like BERT) adapt based on sentence usage.
- Conclusion: Best at capturing semantics.

Q3

Logistic Regression performed slightly better than Naive Bayes, with an accuracy of 86% vs. 80% and a higher F1-score. This could be because Logistic Regression can better capture complex decision boundaries, especially when text features have overlapping distributions.

Q4

The passage begins with a neutral tone, building tension in the middle segments as conflict arises. Segment 3 shows a strong negative dip, possibly reflecting fear or sadness. The emotional tone recovers slightly by segment 5, ending on a hopeful or resolved note.

Q5

1. Why is *lemmatization* often preferred over *stemming*?

Lemmatization returns the base or dictionary form of words using vocabulary and grammar, making it more accurate than stemming, which may cut words crudely.

2. How does *TF-IDF* down-weight common words?

TF-IDF reduces the weight of frequently occurring words across documents by using the inverse document frequency, which penalizes terms that appear in many documents.

3. Describe the *curse of dimensionality* in text data.

In high-dimensional text representations like BoW or TF-IDF, models struggle due to sparse features, leading to overfitting and poor generalization.

4. When should you use *word embeddings* instead of BoW/TF-IDF?

Use word embeddings when you need to capture semantic relationships between words or reduce dimensionality for deep learning models.

5. How can *POS tagging* enhance NLP pipelines?

POS tagging helps identify the grammatical role of words, improving tasks like parsing, named entity recognition, and lemmatization accuracy.