



# **NLP FINAL PROJECT**

MADANI - MAKKI QURAN VERSES  
CLASSIFICATION

---



done for :  
DIA ABUZENEH

done by :  
MONJED TAMIMI 211069

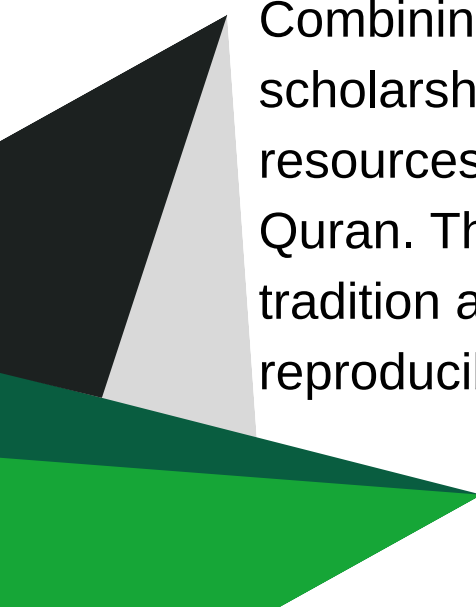




## About the project:

The Quran, Islam's holy book, consists of verses classified as Makkie or Madannie based on where they were revealed. Makkie verses from Mecca emphasize spiritual and ethical teachings, while Madannie verses from Medina address social, legal, and communal issues. Proper classification is crucial for understanding their historical and theological context.

Traditionally, this classification relies on scholars' expertise and historical knowledge, which is time-consuming. Our project aims to streamline and enhance this process using modern machine learning techniques. By employing algorithms like Naive Bayes, Support Vector Machines (SVM), Logistic Regression, and Cosine Similarity, we seek to automate the classification of Quranic verses.



Combining computational methods with traditional scholarship supports scholars, improves educational resources, and deepens our understanding of the Quran. This project exemplifies the integration of tradition and technology, providing a systematic and reproducible approach to studying the Quranic text.

## Accuracies for the models

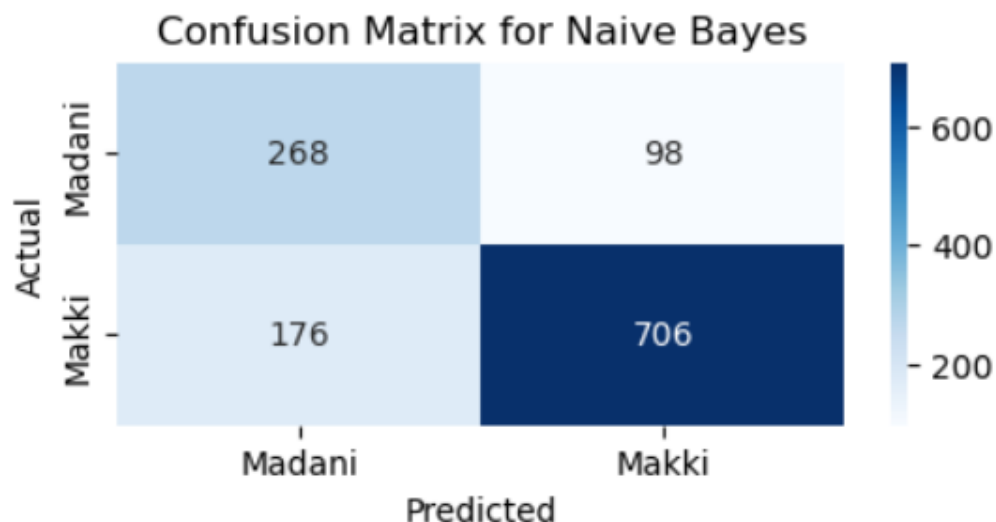
The accuracy of the Naive Bayes classifier is: 78.04%

The accuracy of the SVM classifier is: 73.32%

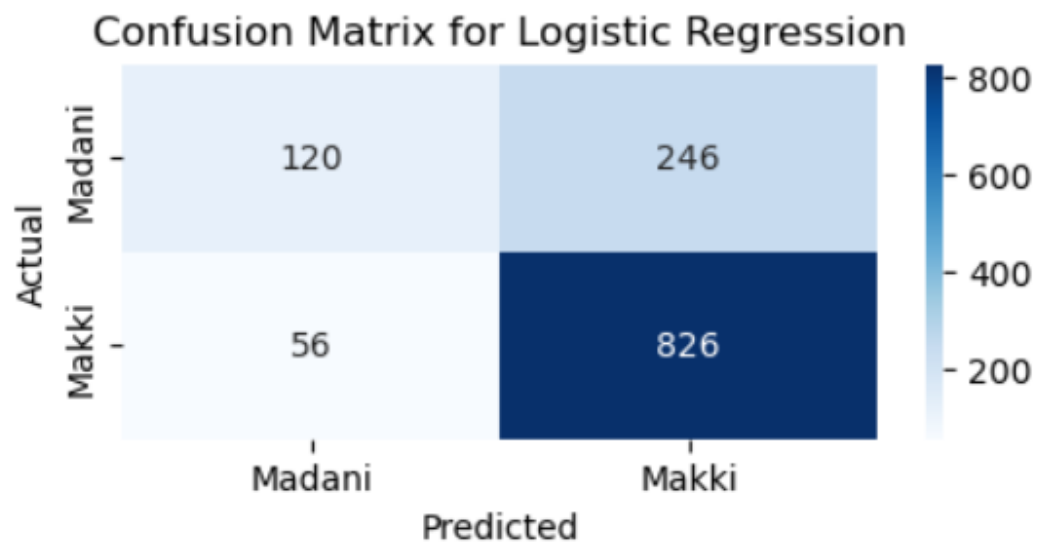
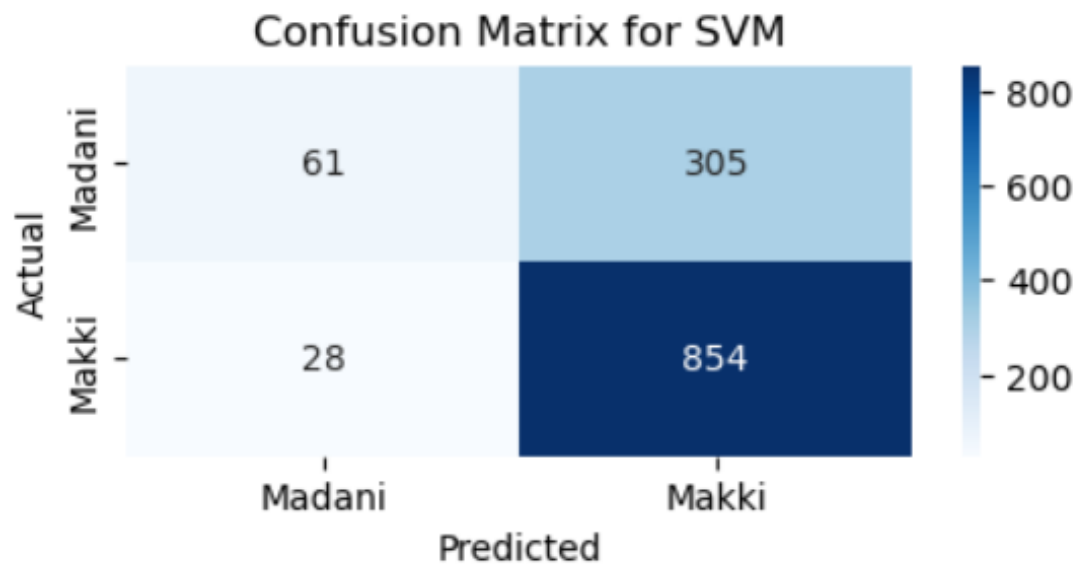
The accuracy of the Logistic Regression classifier is: 75.80%

The accuracy of the Cosine Similarity model is: 70.67%

## Confusion matrix for each model

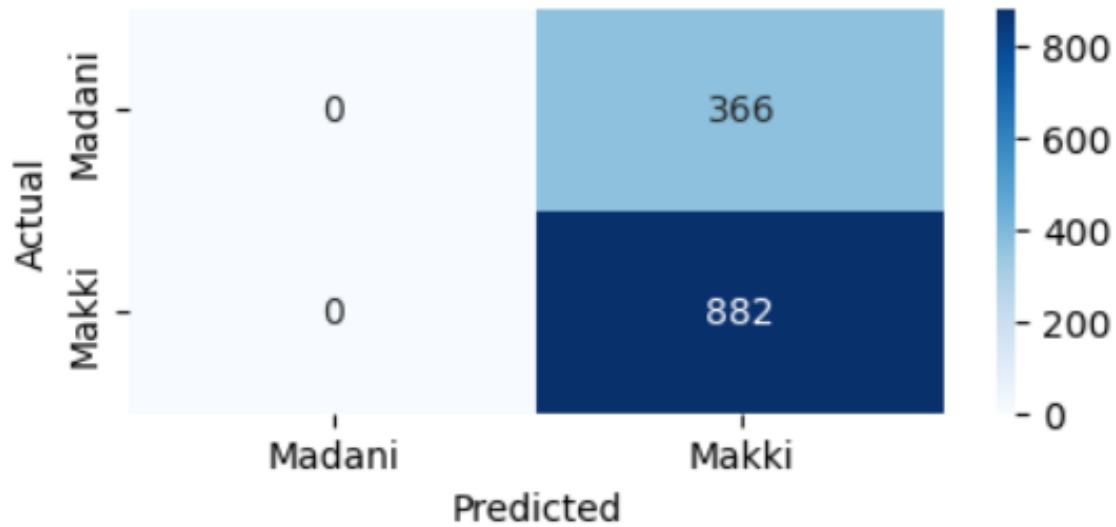


03



04

Confusion Matrix for Cosine Similarity Model



## Analysis of Misclassified Cases

### Analysis of Misclassified Cases by Naive Bayes:

Case 1:

Verse: ولا تجادل عن الذين يختانون أنفسهم<sup>٢</sup> إن الله لا يحب من كان خوانا أثيما

True Label: Madani

Predicted Label: Makki

Case 2:

Verse: لولا إذ سمعتموه ظن المؤمنون والمؤمنات بأنفسهم خيرا وقالوا هذا إفك مبين

True Label: Madani

Predicted Label: Makki

Case 3:

Verse: وقال موسى يا فرعون إني رسول من رب العالمين

True Label: Makki

Predicted Label: Madani

Case 4:

Verse: من كان يريد ثواب الدنيا فعند الله ثواب الدنيا والآخرة وكان الله سميعا بصيرا

True Label: Madani

Predicted Label: Makki

Case 5:

Verse: وما أرسلنا قبلك من المرسلين إلا إنهم ليأكلون الطعام ويمشون في الأسواق

وجعلنا بعضكم لبعض فتنة أتصبرون وكان ربك بصيرا

True Label: Makki

Predicted Label: Madani

## Reasons for Misclassification by Naive Bayes:

Possible reasons include insufficient word features, similar context across categories, or overlap in linguistic styles between Makki and Madani verses. Naive Bayes assumes that features are independent, which is not always the case with textual data, resulting in errors. Additionally, the model's performance can be affected by the quantity and quality of the training data, leading to potential misclassifications.

## Analysis of Misclassified Cases by SVM:

Case 1:

Verse: ولا تجادل عن الذين يختانون أنفسهم إن الله لا يحب من كان خوانا أثيما

True Label: Madani

Predicted Label: Makki

Case 2:

Verse: ولله ما في السماوات وما في الأرض وإلى الله ترجع الأمور

True Label: Madani

Predicted Label: Makki

Case 3:

Verse: لولا إذ سمعتموه ظن المؤمنون والمؤمنات بأنفسهم خيرا وقالوا هذا إفك مبين

True Label: Madani

Predicted Label: Makki

Case 4:

Verse: يَمْنُونَ عَلَيْكَ أَنْ أَسْلَمُوا قُلْ لَا تَمْنُوا عَلَيَّ إِسْلَامَكُمْ بَلِ اللَّهُ يَمُنْ عَلَيْكُمْ أَنْ هَدَاكُمْ  
لِلْإِيمَانِ إِنْ كُنْتُمْ صَادِقِينَ

True Label: Madani

Case 5:

Verse: إِنْ الَّذِينَ كَفَرُوا يَنْفِقُونَ أَمْوَالَهُمْ لِيَصُدَّوْا عَنْ سَبِيلِ اللَّهِ فَسَيَنْفِقُونَهَا ثُمَّ تَكُونُ عَلَيْهِمْ  
حَسْرَةً ثُمَّ يُغْلَبُونَ وَالَّذِينَ كَفَرُوا إِلَىٰ جَهَنَّمَ يُحْشَرُونَ

True Label: Madani

Predicted Label: Makki

## SVM Misclassification Reasons

Possible reasons include insufficient word features, similar context across categories, or overlap in linguistic styles between Makki and Madani verses. SVMs can be sensitive to the chosen kernel and hyperparameters, which might not always capture the complexity of the data. Additionally, the model's performance can be affected by the quantity and quality of the training data, leading to potential misclassifications.

## Analysis of Misclassified Cases by Logistic Regression:

Case 1:

Verse: وَلَا تَجَادَلْ عَنْ الَّذِينَ يُخْتَانُونَ أَنْفُسَهُمْ إِنَّ اللَّهَ لَا يُحِبُّ مَنْ كَانَ خَوَانًا أَثِيمًا

True Label: Madani

Predicted Label: Makki

Case 2:

Verse: وَلِلَّهِ مَا فِي السَّمَاوَاتِ وَمَا فِي الْأَرْضِ وَإِلَى اللَّهِ تُرْجَعُ الْأُمُورُ

True Label: Madani

Predicted Label: Makki

Case 3:

Verse: لَوْلَا إِذْ سَمِعْتُمُوهُ ظَنَّ الْمُؤْمِنُونَ وَالْمُؤْمِنَاتُ بِأَنْفُسِهِمْ خَيْرًا وَقَالُوا هَذَا إِفْكٌ مُبِينٌ

True Label: Madani

Predicted Label: Makki

Case 4:

Verse: يَمْنُونَ عَلَيْكَ أَنْ أَسْلَمُوا قُلْ لَا تَمْنُوا عَلَيَّ إِلَّا سَلَامُكُمْ ۚ بَلِ اللَّهُ يَمُنُ عَلَيْكُمْ أَنْ هَدَاكُمْ

لِلْإِيمَانِ إِنْ كُنْتُمْ صَادِقِينَ

True Label: Madani

Predicted Label: Makki

Case 5:

Verse: سَمَاعُونَ لِلْكَذِبِ أَكَالُونَ لِلْسَحْتِ ۖ إِنْ جَاءُوكَ فَاحْكُم بَيْنَهُمْ أَوْ أَعْرِضْ عَنْهُمْ ۚ وَإِنْ

تَعْرِضْ عَنْهُمْ فَلَنْ يَضُرُّوكَ شَيْئًا ۚ وَإِنْ حَكَمْتَ فَاحْكُم بَيْنَهُم بِالْقِسْطِ ۚ إِنَّ اللَّهَ يُحِبُّ الْمُقْسِطِينَ

True Label: Madani

Predicted Label: Makki

## Logistic Regression Misclassification Reasons

Possible reasons include insufficient word features, similar context across categories, or overlap in linguistic styles between Makki and Madani verses. Logistic Regression assumes a linear relationship between features and the target, which might not always hold true for textual data. Furthermore, the model's performance can be impacted by the quantity and quality of the training data, leading to potential misclassifications.

## Analysis of Misclassified Cases by Cosine Similarity Model:

Case 1:

Verse: وَلَا تَجَادِلْ عَنِ الَّذِينَ يَخْتَانُونَ أَنْفُسَهُمْ ۚ إِنَّ اللَّهَ لَا يُحِبُّ مَنْ كَانَ خَوَانًا أَثِيمًا

True Label: Madani

Predicted Label: Makki

Case 2:

Verse: وَلِلَّهِ مَا فِي السَّمَاوَاتِ وَمَا فِي الْأَرْضِ ۚ وَإِلَى اللَّهِ تُرْجَعُ الْأُمُورُ

True Label: Madani

Predicted Label: Makki

Case 3:

Verse: لَوْلَا إِذْ سَمِعْتُمُوهُ ظَنَّ الْمُؤْمِنُونَ وَالْمُؤْمِنَاتُ بِأَنْفُسِهِمْ خَيْرًا وَقَالُوا هَذَا إِفْكٌ مُبِينٌ

True Label: Madani

Predicted Label: Makki



Case 4:

Verse: يَمْنُونَ عَلَيْكَ أَنْ أَسْلَمُوا قُلْ لَا تَمْنُوا عَلَيَّ إِسْلَامَكُمْ بَلِ اللَّهُ يَمُنْ عَلَيْكُمْ أَنْ هَدَاكُمْ  
لِلْإِيمَانِ إِنْ كُنْتُمْ صَادِقِينَ

True Label: Madani

Predicted Label: Makki

Case 5:

Verse: وَالْوَالِدَاتُ يُرْضِعْنَ أَوْلَادَهُنَّ حَوْلِينَ كَامِلِينَ ۖ لِمَنْ أَرَادَ أَنْ يُتِمَّ الرَّضَاعَةَ ۚ وَعَلَى الْمَوْلُودِ لَهُ  
لَهُ رِزْقُهُنَّ وَكِسْوَتُهُنَّ بِالْمَعْرُوفِ ۚ لَا تُكَلِّفُ نَفْسٌ إِلَّا وُسْعَهَا ۚ لَا تَضَارُّ وَالِدَةُ بَوْلِهَا وَلَا مَوْلُودٌ لَهُ  
بَوْلُهُ ۚ وَعَلَى الْوَارِثِ مِثْلُ ذَلِكَ ۚ فَإِنْ أَرَادَا فِصَالًا عَنْ تَرَاضٍ مِنْهُمَا وَتَشَاوُرٍ فَلَا جُنَاحَ عَلَيْهِمَا ۚ  
وَإِنْ أَرَدْتُمْ أَنْ تُسْتَرْضِعُوا أَوْلَادَكُمْ فَلَا جُنَاحَ عَلَيْكُمْ إِذَا سَلَّمْتُمْ مَا آتَيْتُم بِالْمَعْرُوفِ ۚ وَاتَّقُوا اللَّهَ  
وَاعْلَمُوا أَنَّ اللَّهَ بِمَا تَعْمَلُونَ بَصِيرٌ

True Label: Madani

Predicted Label: Makki

## Cosine Similarity Misclassification Reasons

Possible reasons include insufficient word features, similar context across categories, or overlap in linguistic styles between Makki and Madani verses. The chosen threshold for cosine similarity might not be optimal, leading to incorrect classifications. Moreover, the performance can be affected by the quality of the vectorized text data and the inherent limitations of using similarity measures alone for classification.

### In conclusion :

we cant reallyly determine and know exactly why there is missclassification because we dont know what the model used exactly as reference, the hardet thing is to know a verse is Madani or Makki is depending on the context and when it was first released...etc

**full code in python3**