

גירסה 17.8

‏24/10/2021

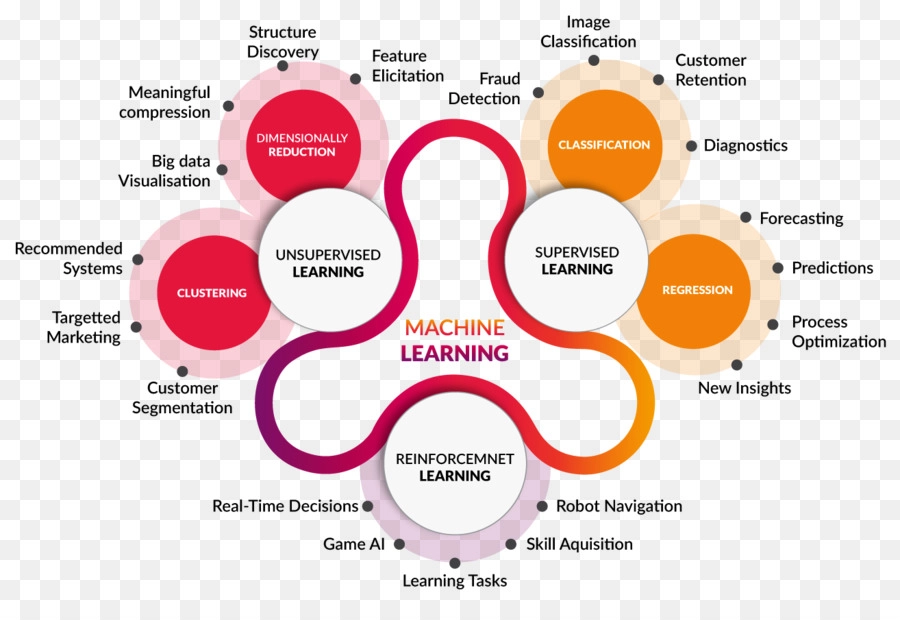
Using machine learning algorithms to classify the Hamzat Alwasl and Hamzat Alqatae

Solving the problem of the wrong use of the hamzat "Wasl" and "Gtaa" in Arabic texts

קוד פרויקט: לדוגמא 21220570

דוח מכין - פרויקט גמר תשפ"ב

מנחה אקדמי 1: מר יורם סגל



מגיש

שם סטודנט א: 025879123 נזאר מערוף

חתימת מנחים

תוכן עניינים

[1 תקציר באנגלית 2](#_Toc85816129)

[2 מבוא 3](#_Toc85816130)

[2.1 כיצד קשור או משתלב הפרויקט לתחום כללי כלשהו 4](#_Toc85816131)

[2.2 הגדרת הבעיה 4](#_Toc85816132)

[2.3 האתגר הטכנולוגי 5](#_Toc85816133)

[3 דרכי פתרון הבעיה 5](#_Toc85816134)

[4 תוצר מצופה מהפרויקט 6](#_Toc85816135)

[תיאור רעיון דומה שיכול להוות השראה 9](#_Toc85816136)

[5 סיכונים, אי וודאות ואילוצי הפרויקט 9](#_Toc85816137)

[6 מקורות קריאה (זה רק דוגמה) 10](#_Toc85816138)

[רשימת נספחים 11](#_Toc85816139)

[נספח א - סכמת בלוקים 12](#_Toc85816140)

[נספח ב - טבלת אבני דרך ותוצרים 13](#_Toc85816141)

[נספח ג - טבלת משימות (לא חובה במסגרת דוח מכין - לפי החלטת המנחה) 14](#_Toc85816142)

[נספח ד 15](#_Toc85816143)

[נספח ה - גאנט (אופציונלי בהתאם לשיקול דעתו של המנחה) 16](#_Toc85816144)

# תקציר באנגלית

The Arabic language as considered a living organism that grows and develops through the correct practice and application of all it literature and it syntactic, morphological, semantic, and lexical branches. In this research study, we review the contribution of technology to the development of Arabic language, especially the correct writing of Hamzat Alwasl and Hamzat Alqatae. This project aims to building a smart Classifier model to classifying Arabic words beginning with the letter Alif and this Hamzat Alwasl and Hamzat Alqatae through the using of artificial intelligence techniques in several and machine learning algorithms to establish accurate and correct criteria in writing the Hamzat Alwasl and Hamzat Alqatae correctly.

Consequently, technology would be adapted to contributes to the service of Arabic language.

This research relied on the compilation of Arabic words beginning with hamza by designing a digital questionnaire. The task of this questionnaire is to collect the largest number of words beginning with Hamza and to classifying them as Hamzat Alwasl and Hamzat Alqatae according to the gram The Arabic language is considered a living organism that grows and develops through the correct practice and application of all its literature and its syntactic, morphological, semantic, and lexical branches. In this research study, we review the contribution of technology to the development of Arabic language, especially the correct writing of Hamzat Alwasl and Hamzat Alqatae. This project aims to building a smart Classifier model to classifying Arabic words beginning with the letter Alif and this Hamzat Alwasl and Hamzat Alqatae through the using of artificial intelligence techniques in several and machine learning algorithms to establish accurate and correct criteria in writing the Hamzat Alwasl and Hamzat Alqatae correctly.

Consequently, technology would be adapted to contributes to the service of Arabic language.

This research relied on the compilation of Arabic words beginning with hamza by designing a digital questionnaire. The task of this questionnaire is to collect the largest number of words beginning with Hamza and to classifying them as Hamzat Alwasl and Hamzat Alqatae according to the grammatical rules followed in this process. The questionnaire was circulated in the internet web and was filled out by fifty specialists in syntax with different academic ranks. The total number of classified words reached 400 words, and after processing and excluding the repeated words, 101 words, we obtained 299 valid words to be applied to the Classifier model, and based on the size and type of the collected data and the classification mechanism followed, classification algorithms were applied that fit the collected sample such as the vector support machine algorithms , Naif Biz algorithms, and Nearest neighbor algorithms by using Python language and sk-learn library.

After training the used classificatory models and Measuring the accuracy of the algorithms, it was quite apparent that the vector support machine (SVM) algorithms had obtained the highest accuracy of 92%, which is high and sufficient to solve the problem of the project.

matical rules followed in this process. The questionnaire was circulated in the internet web and was filled out by fifty specialists in syntax with different academic ranks. The total number of classified words reached 400 words, and after processing and excluding the repeated words, 101 words, we obtained 299 valid words to be applied to the Classifier model, and based on the size and type of the collected data and the classification mechanism followed, classification algorithms were applied that fit the collected sample such as the vector support machine algorithms , Naif Biz algorithms, and Nearest neighbor algorithms by using Python language and sk-learn library.

After training the used classificatory models and Measuring the accuracy of the algorithms, it was quite apparent that the vector support machine (SVM) algorithms had obtained the highest accuracy of 92%, which is high and sufficient to solve the problem of the project.

# מבוא

We can define [machine learning](https://www.zdnet.com/article/what-is-machine-learning-everything-you-need-to-know/) as following the science of [artificial intelligence](https://www.zdnet.com/article/what-is-ai-heres-everything-you-need-to-know-about-artificial-intelligence/#link=%7B%22role%22:%22standard%22,%22href%22:%22https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/%22,%22target%22:%22_blank%22,%22absolute%22:%22%22,%22linkText%22:%22artificial%20intelligence%22%7D) that is concerned with designing algorithms and techniques that allow computers to process self-learning with the possibility of developing these algorithms.

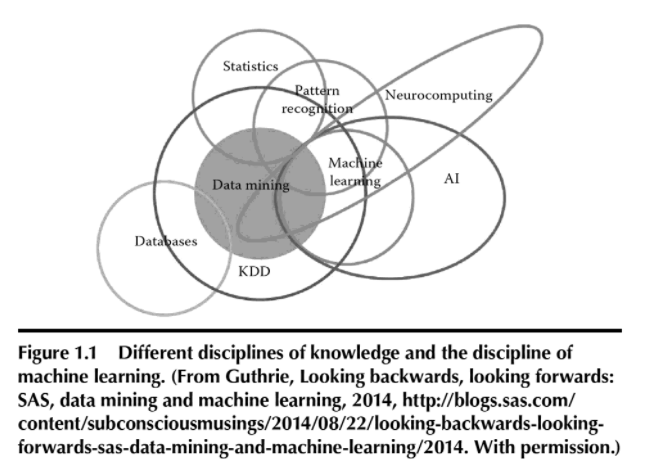
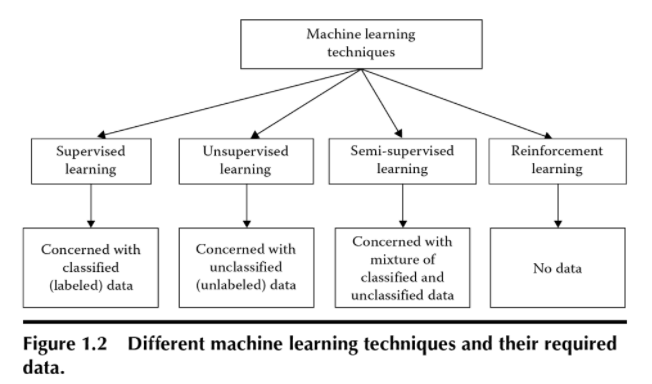
We can review two types of machine learning: inductive learning and deductive learning, inductive learning is known as inductive learning and works on the principle of also deducing general rules from data, and this differs from inductive learning where the learner is given the rules that he needs to apply.

Machine learning depends on the principle of systems learning from available data, identifying appropriate patterns and making decisions without human intervention, and since the primary task of machine learning is to extract information from data, machine learning is a method of data analysis that automates building data analysis models known as analysis models. Predictive, this pattern is known as the predictive or taxonomic analysis model, these models allow researchers, data scientists, and analysts to make decisions that enable reliable results.

Building a data analysis model is subject to several mechanisms or ways to reach accurate results, as the model building stage requires the implementation of several arranged and sequential mathematical or logical steps to reach a trained analytical model that can predict values ​​closer to the truth or accurately classify the data. This group of steps is called the name Algorithm, a data classification model is also called a classifier. In this project, I will address the use of [machine learning algorithms](https://d1wqtxts1xzle7.cloudfront.net/38147943/jait0101-with-cover-page-v2.pdf?Expires=1634839344&Signature=aUPJxazqNewl77PCgq1oVP~bM4Jsi5B60rmprhEaE1F5opqtdaSnUnMRbdcRiYDbGSqsAV45PsFpUZV0EYyA~j2t8pfZqO0iwnXVoozc9Rt1RQrVZri-ia~OcjwR9dEfdjWjJhpzyIC-QfPMYg7nn5ScNMbmgpzmvB4QmeMHc~5kXGHS9b7EF9zinvZjtcpLuHG8O1Jir2Fyf1mzzQF-C1n3EPIqlnI4~1RDoIaDpdrQmTyoWmyqz15FWRYz9xo8hNe64L2nTJo6HvgW0oc9JvLrYMAUo80MA8guRmnckgZkozdQI244BnW3K45XCQ42O88ZkcTL7b9gEel2fmoihg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA#page=6) models in classifying the Hamzat Alwasl and Hamzat Alqatae, and these algorithms will be applied to Arabic words because the Arabic language has not been as well studied and researched as other languages.

These are benefits of machine learning and why research in machine learning is now what could not be avoided or neglected. Using machine learning techniques make life easier for computer users. These are the importance of machine learning:

* Some tasks cannot be defined well expect by example, that is we might be able to specify input and output pairs but not a concise relationship between inputs and desired outputs. We would like machine to be able to adjust their internal structure to produce correct outputs for many sample inputs and thus suitably constrain their input and output function to approximate the relationship implicit in the examples.
* Human designer often produce machine that do not work as well as desired in the environments in which they are used. In fact, certain characteristics of the working environment might not be completely known at design time. Machine learning methods can be used for on the job improvement of existing machine designs.



## כיצד קשור או משתלב הפרויקט לתחום כללי כלשהו

We can define [machine learning](https://www.zdnet.com/article/what-is-machine-learning-everything-you-need-to-know/) as following the science of [artificial intelligence](https://www.zdnet.com/article/what-is-ai-heres-everything-you-need-to-know-about-artificial-intelligence/#link=%7B%22role%22:%22standard%22,%22href%22:%22https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/%22,%22target%22:%22_blank%22,%22absolute%22:%22%22,%22linkText%22:%22artificial%20intelligence%22%7D) that is concerned with designing algorithms and techniques that allow computers to process self-learning with the possibility of developing these algorithms.

We can review two types of machine learning: inductive learning and deductive learning, inductive learning is known as inductive learning and works on the principle of also deducing general rules from data, and this differs from inductive learning where the learner is given the rules that he needs to apply.

Machine learning depends on the principle of systems learning from available data, identifying appropriate patterns and making decisions without human intervention, and since the primary task of machine learning is to extract information from data, machine learning is a method of data analysis that automates building data analysis models known as analysis models. Predictive, this pattern is known as the predictive or taxonomic analysis model, these models allow researchers, data scientists, and analysts to make decisions that enable reliable results.

Building a data analysis model is subject to several mechanisms or ways to reach accurate results, as the model building stage requires the implementation of several arranged and sequential mathematical or logical steps to reach a trained analytical model that can predict values ​​closer to the truth or accurately classify the data. This group of steps is called the name Algorithm, a data classification model is also called a classifier. In this project, I will address the use of [machine learning algorithms](https://d1wqtxts1xzle7.cloudfront.net/38147943/jait0101-with-cover-page-v2.pdf?Expires=1634839344&Signature=aUPJxazqNewl77PCgq1oVP~bM4Jsi5B60rmprhEaE1F5opqtdaSnUnMRbdcRiYDbGSqsAV45PsFpUZV0EYyA~j2t8pfZqO0iwnXVoozc9Rt1RQrVZri-ia~OcjwR9dEfdjWjJhpzyIC-QfPMYg7nn5ScNMbmgpzmvB4QmeMHc~5kXGHS9b7EF9zinvZjtcpLuHG8O1Jir2Fyf1mzzQF-C1n3EPIqlnI4~1RDoIaDpdrQmTyoWmyqz15FWRYz9xo8hNe64L2nTJo6HvgW0oc9JvLrYMAUo80MA8guRmnckgZkozdQI244BnW3K45XCQ42O88ZkcTL7b9gEel2fmoihg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA#page=6) models in classifying the Hamzat Alwasl and Hamzat Alqatae, and these algorithms will be applied to Arabic words because the Arabic language has not been as well studied and researched as other languages.

**Key words:** Artificial Intelligent, Machine Learning Algorithms, Hamza, Wasl, Gtaa, Classification Algorithms, Arabic Language.

## הגדרת הבעיה

The main problem of the project lies in the wrong use of the Hamzat Alwasl and Hamzat Alqatae in the wrong place, especially when writing words beginning with the letter "Alif", as many make mistakes in drawing the hamza or dropping it, which weakens the strength of the language. This project contributes to the development of the Arabic language, by building a smart classification model for editors, practitioners, and all language users from the correct classification of the Hamzat Alwasl and Hamzat Alqatae. This project seeks to answer the following questions in light of making use of machine learning algorithms used for classification and prediction:

* How can we distinguish between the Hamzat Alwasl and Hamzat Alqatae in Arabic texts?
* What are the benefits that you reaped when classifying the Hamzat Alwasl and Hamzat Alqatae in Arabic texts?
* What are the appropriate machine learning algorithms for classifying the Hamzat Alwasl and Hamzat Alqatae in Arabic texts?
* Solving the problem of the wrong use of the hamzat and wasl in Arabic texts.
* Identifying the most appropriate classification algorithms with in machine learning algorithms to determine the correct positions, using the Hamzat Alwasl and Hamzat Alqatae at the beginning of a words.
* Measuring the quality of the most famous classification algorithms in distinguishing between the hyphens and hyphens at the beginning of a words.

## האתגר הטכנולוגי

This study is one of the new studies in the field of grammatical and technical studies dealing with the use of machine learning algorithms to classify the Hamzat Alwasl and Hamzat Alqatae in Arabic texts, and as far as I know, and through my research, I have not found any studies that have been exposed to this subject before because of the difficulty of using software libraries in dealing with Arabic texts deservedly, but there is a great similarity and overlap of other studies associated with the project to use these algorithms in similar and interesting applications, and I have benefited from them, and helped me enter my project, including:

* Exploration of opinions in arabic comparison sentences [**[1]**](https://repository.nauss.edu.sa/bitstream/handle/123456789/53767/%D8%AA%D9%86%D9%82%D9%8A%D8%A8%20%D8%A7%D9%84%D8%A2%D8%B1%D8%A7%D8%A1%20%D9%81%D9%8A%20%D8%AC%D9%85%D9%84%20%D8%A7%D9%84%D9%85%D9%82%D8%A7%D8%B1%D9%86%D8%A9%20%D8%A7%D9%84%D8%B9%D8%B1%D8%A8%D9%8A%D8%A9.pdf?sequence=1&isAllowed=y)
* Comparative study of opinion exploration algorithms and analysis of emotions and theirapplications[**[2]**](https://www.iasj.net/iasj/download/0809c21eebbee97f)

Bottom line, we can use the new model as a software tool that can be combined with all software and technical systems that display readable or written Arabic text using software systems to categorize the Hamza in words that are started with a Hamzat Alwasl and Hamzat Alqatae.

# דרכי פתרון הבעיה

In this part of the project[**[3]**](https://www.iasj.net/iasj/download/0809c21eebbee97f), I will address the first four stages of the application of classifier models, and the remaining two stages will be addressed in the third part of the project.

* The stage of designing the questionnaire and describing the data:

The Arabic words beginning with hamza were collected by designing a digital questionnaire (see Appendix A). Specialists in the Arabic language, and the specialists have classified it as two hyphenated or suffixes according to well-known grammatical rules

The words extracted from the sample texts were classified into words that start with a hamzat Alwasl referred to in this project as (Wasl) and words that start with a hamzat Alqatae referred to in this project as (Gtaa).

As for the independent variables (characteristics), they were divided into three properties according to the grammatical rules.

The format of the dataset file was chosen from type (csv), and it contained the properties (independent variables).

* Data cleaning stage:

Certainly, when the quality of the data is low, this will inevitably affect the results of the analysis.

After performing data cleaning operations on the collected texts, and after processing and excluding duplicate words, as well as completing the missing data, the number of excluded words became 101, we got 299 valid words to apply to the workbook model.

* Data coding and representation:
* The stage of building and training classifier models:

The represented data was divided into training data and experimental data with a percentage out of the total data recorded as a prelude to building a training data model using the selected classification algorithms, and it was selected in proportion to fit the sample size and data values with its algorithms.

# תוצר מצופה מהפרויקט

The data in today's is growing day by day[**[4]**](https://www.naya-college.co.il/%D7%91%D7%99%D7%92-%D7%93%D7%90%D7%98%D7%94/) and its sources are multiple, and this leads to the exposure of this data to many problems that reduce the quality of the data such as the large number of missing data and the inconsistency of the data, so in this project I divided the stages of applying data classifiers models into six stages[**[5]**](https://www.iasj.net/iasj/download/0809c21eebbee97f) that started At the stage of designing the questionnaire and collecting data, and ending with the stage of measuring the accuracy of classifier models, as shown in the following figure (1):

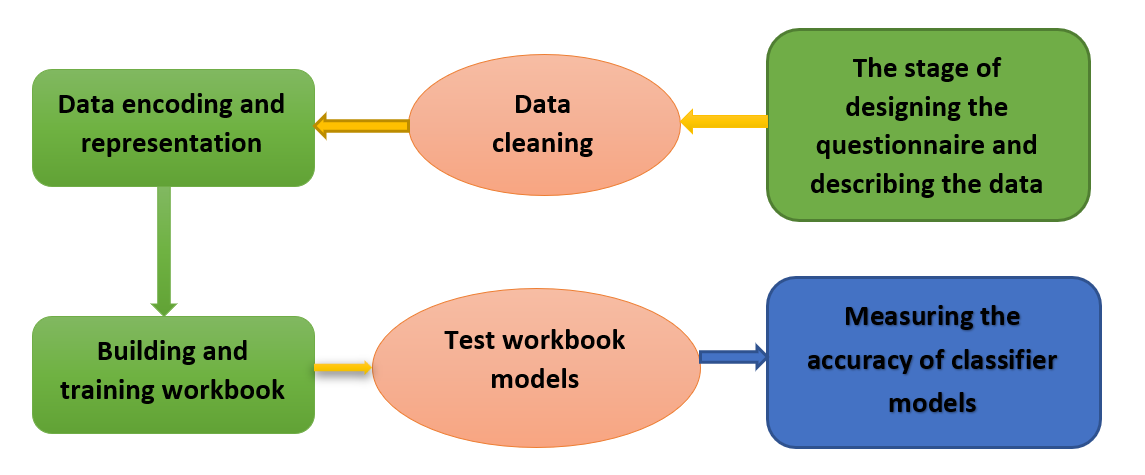


Figure (1): Standard block diagram.

**First: Test the models of the workbooks**

After completing the stage of building classifier models using (NB, SVM KNN) algorithms, we moved to the stage of testing these models so that we can measure their quality later. The models were tested by making predictions on the models after training them using (sklearn) library functions in python language.

**Second: Measuring the accuracy of classifier models**

The stage of evaluating the results of data mining models is one of the important stages that enables us to define the most effective model. The effectiveness of the model is measured through the accuracy of the applicable plan. The nature of the data used in building models plays a key role in their effectiveness. There are many statistical methods that test classification models. We show the most important of them as follows:

* **Calculation of average accuracy**

It is the arithmetic average of the ratios of the accuracy of the correct predictions for each category provided by the model to the number of actual classifications for this category in the test data set. A percentage when calculating the average accuracy, as its accuracy reached (94%), while the (NB) and. algorithms obtained (KNN) on average accuracy (54%) and (85%), respectively.

* **Calculation of total accuracy**

It is the ratio of the sum of correct predictions given by the model to the sum of the actual ratings in the test data set.

* **Finding the confusion matrix**

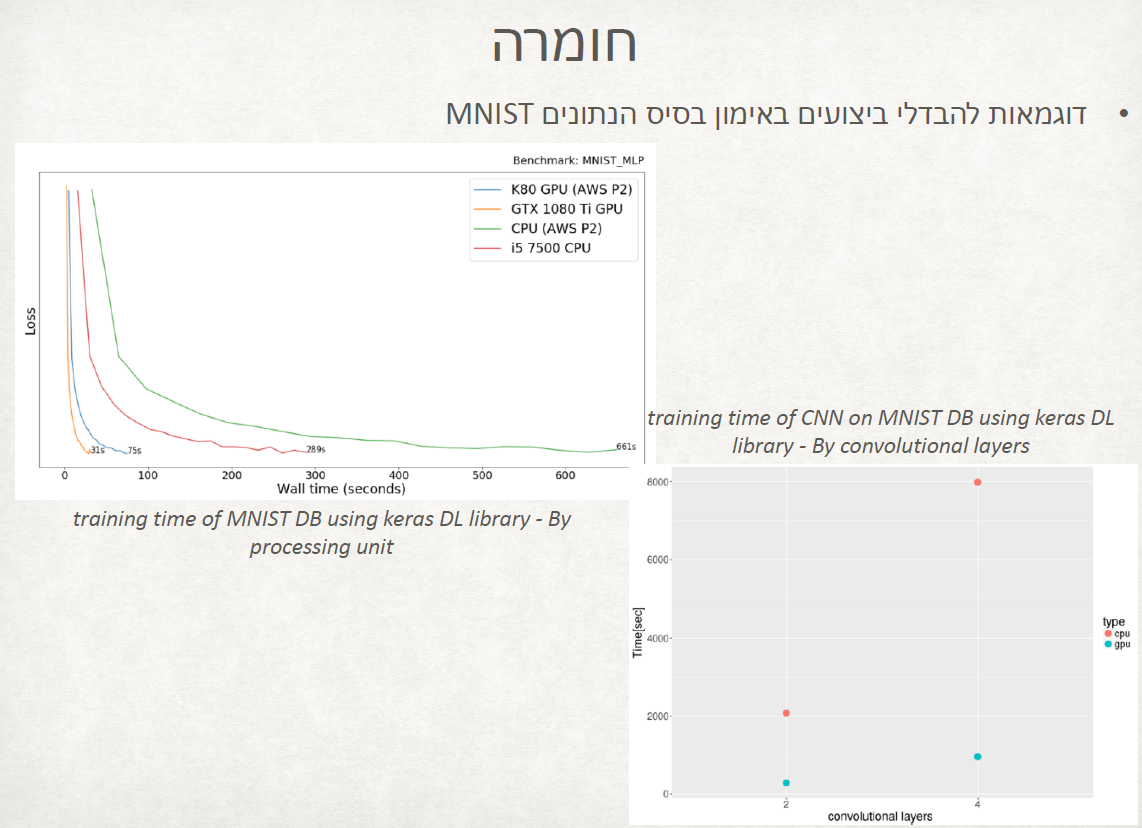
The confusion matrix shows the number of correctly predicted cases and the number of incorrectly predicted cases in the test set for each item compared to the number of actual cases for those items. The rank of this matrix is n × n where n is the number of items in the column of the target variable (the dependent variable), and the confusion matrix was found for the algorithms used to classify the dependent variable Outcome) into two classes (the Wasl link and the cut link (Gtaa).

**Third: the results**

We have used the Support Vector Machine (SVM) algorithm, the Knife Base algorithm (NB) and the Nearest Neighbour algorithm (KNN) to classify the hyphens and slashes (at the beginning of the word). Based on the classification process, a training model was made for all these algorithms, and it became clear by calculating the accuracy of the model for each algorithm that it exceeded the SVM algorithm, as it obtained the highest accuracy rate of the model amounting to 92%, which is high and sufficient to solve the basic research problem. The main reason for the algorithm to obtain such a high percentage was the large number of characteristics used in the classification (8 independent variables were used), which contributed to making the classification model more realistic.

**Fourth: Research outcomes and desired benefits**

The aim of this completed project was to build a smart model that categorizes the link and cutting (at the beginning of the word) by using data classification algorithms in order to establish accurate and correct criteria in accurate writing Arabic texts to contribute and help adapt the technology to measure the quality of the most famous classification algorithms in distinction to serve the Arabic language, as well as the goal of research between the link and cutting whispers at the beginning of the word, and this intelligent model that was designed can contribute to the development of the most famous classification algorithms in the service of the Arabic language, as well as the goal of research between the two link and cutting whispers at the beginning of the word, and this intelligent model that was designed can contribute to the development of the most famous classification algorithms in the service of the Arabic language. Arabic as follows:

* The new model can be used to review scientific research and ensure that the whisper is drawn in words correctly, contributing to the completeness of the elements of scientific research.
* The new model can be used to review news articles published on various social media by installing it in the browser plugins section to make sure that the whisper is drawn correctly in all the words of the published articles, which contributes to the development of the new Arab media.
* The new model can be used with tour applications and considered as an authentic part of mobile operating systems to make sure that translated commands are written in Arabic and started with a whisper correctly.
* The new model can be matched with translation systems used in conferences, aircraft and trains, helping to display the humming properly.
* **Tools and obedience**
  + **Software:**
* Python programming in JetBrains PyCharm development environment.
  + **Useful in libraries:**
* TensorFlow של גוגל
* pandas
* NumPy
* Sklearn
* תמונה שמכילה שולחן

  התיאור נוצר באופן אוטומטיUse of databases

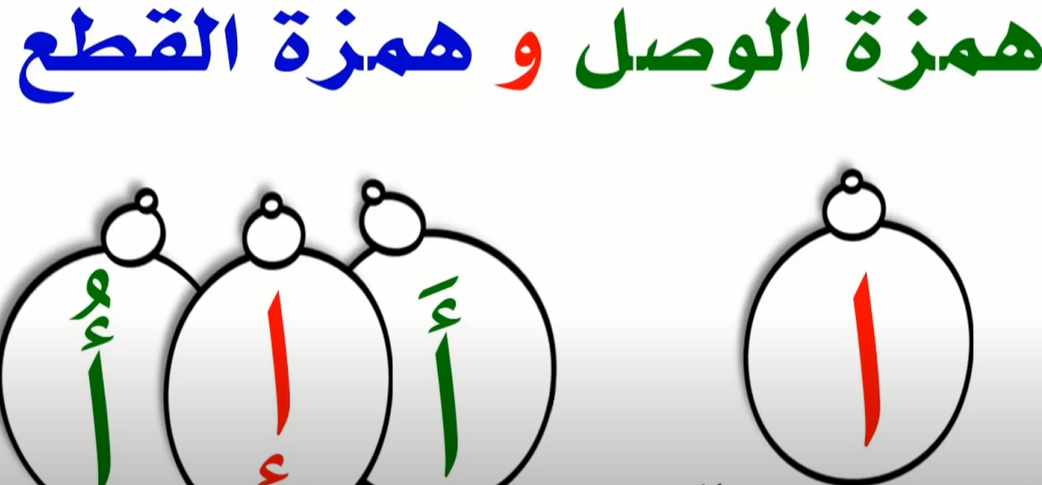
# תיאור רעיון דומה שיכול להוות השראה

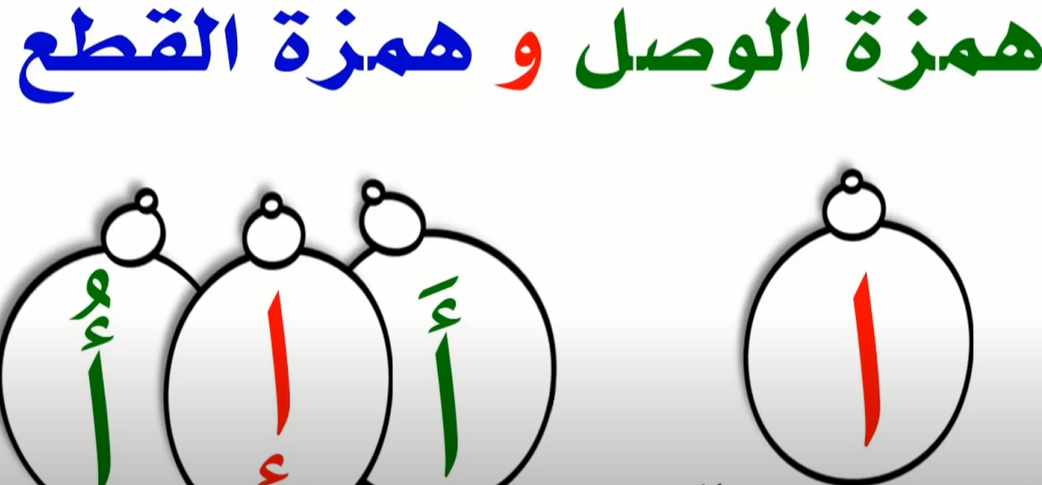
The main problem of the project lies in the incorrect use of the two hamzas in the wrong place, especially when writing words beginning with the letter alif, as many make mistakes in drawing the hamza or dropping it, which weakens the strength of the language.

I got this inspiration (how many people make the mistake of writing the Hamzat Alwasl and Hamzat Alqatae) from videos I watched while trying to help my daughter who is in ninth grade how to know how to choose Hamzat Alwasl or Hamzat Alqatae to the word beginning of the line in Arabic texts.

Funny I recognized that I myself until now would have written them incorrectly.

Attached is a link to the videos:

[](https://www.youtube.com/watch?v=mK8PcK5JpzQ)

[](https://www.youtube.com/watch?v=WLLyn3VWPrM)

# סיכונים, אי וודאות ואילוצי הפרויקט

The main problem of the project lies in the erroneous use of misplaced Hamzat Alwasl and Hamzat Alqatae when writing words with the letter "Alef" in particular, where many mistake the drawing or dropping of the AlHamza, which weakens the power of the language. This project contributes to the development of the Arabic language, by building a smart classification model for editors and practitioners, and all language users from the correct classification of the Hamzat Alwasl and Hamzat Alqatae. This research seeks to answer the following questions considering the use of machine learning algorithms used in classification and prediction:

* How can we distinguish between the two Hamzat Alwasl and Hamzat Alqatae in Arabic texts?
* What are the benefits of classifying the Hamzat Alwasl and Hamzat Alqatae in Arabic texts?
* What machine learning algorithms are suitable for classifying the Hamzat Alwasl and Hamzat Alqatae in Arabic texts?

This study is one of the new studies in the field of grammatical and technical studies that deals with the use of machine learning algorithms to classify the Hamzat Alwasl and Hamzat Alqatae in Arabic texts aptly.

**Recommendations:**

1. Apply algorithms suitable for dealing with purposes with multiple classification categories using indexing algorithms other than those that improvise purposes with binary categories only.
2. Develop sklearn libraries to fully support Arabic.
3. Develop the model to act as a work for arabic audiotexts.

# מקורות קריאה (זה רק דוגמה)

|  |  |
| --- | --- |
| [1] | A. R. A. Rawan A., S. Muhammed N. and F. Polla, "A comprehensive study on sign languages recognition systems using (SVM, KNN, CNN and ANN)," in *Proceedings of the First International Conference on Data Science (DATA '18)*, Madrid, Spain, 2018. |
| [2] | Y. Liao, A. Vakanski ו M. Xian, “A deep learning framework for assessing physical rehabilitation exercises,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* כרך 28, p. 468–477, 2020. |
| [3] | K. Y. L. L. S. L. Jiefu Zhai, “A low complexity motion compensated frame interpolation method,” ב- *Conference: Circuits and Systems, 2005. ISCAS 2005. IEEE International*, 2005. |
| [4] | W. M. Y. H. Yanli Li, “A Spatial Prediction-Based Motion-Compensated Frame Rate Up-Conversion,” January 2019. |
| [5] | G. Hidalgo, “CMU-Perceptual-Computing-Lab/OpenPose,” GitHub, 19 December 2018. [מקוון]. Available: https://github.com/CMU-Perceptual-Computing-Lab/openpose. |
| [6] | S. R. Machlin, J. Chevan, W. W. Yu ו M. W. Zodet, “Determinants of utilization and expenditures for episodes of ambulatory physical therapy among adults,” *Phys Ther,* כרך 91, מס' 7, p. 1018–1029, Jul. 2011. |
| [7] | T. Simon, H. Joo, I. Matthews ו Y. Sheikh, “Hand keypoint detection in single images using multiview bootstrapping,” *InProceedings of the IEEE conference on Computer Vision and Pattern Recognition,* pp. 1145-1153, 2017. |

נספחים

# רשימת נספחים

נספח א - סכמת בלוקים

נספח ב - טבלת אבני דרך ותוצרים

נספח ג - טבלת משימות (לא חובה במסגרת דוח מכין - לפי החלטת המנחה)

נספח ד- חתימה על מסמכים רלוונטיים של המרכז האקדמי רופין

נספח ה - גאנט (אופציונלי בהתאם לשיקול דעתו של המנחה)

# נספח א - סכמת בלוקים

In this project, I divided the stages of applying data classifier models in to several stages, starting with the stage of designing the questionnaire and collecting data and ending with the stage of Measuring the accuracy of the classifier models. As in the following figure (1)

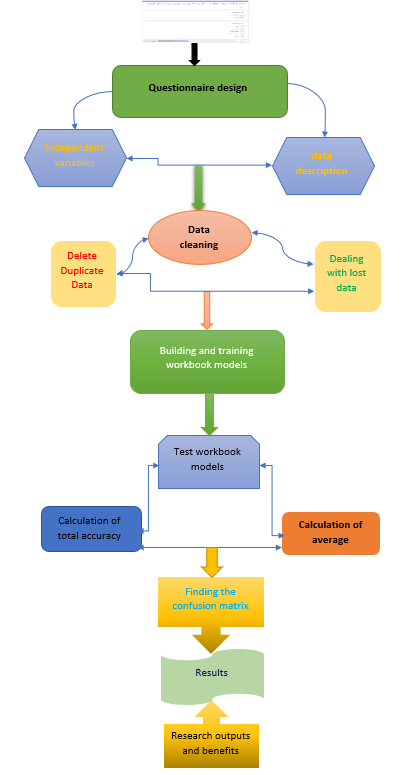


figure (1): Standard block diagram.

# נספח ב - טבלת אבני דרך ותוצרים

מטרת נספח זה היא הוכחת כושר תכנון זמנים עתידי, תחת תנאי אי וודאות.

עליכם להכין במסגרת דוח המכין טבלת אבני דרך נפרדת, המציגה רשימת אבני דרך ממוספרות (7 אבני דרך בדיוק!!!) להלן הפורמט המחייב:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מס' אבן הדרך | תיאור אבן הדרך | תאריך סיום | סה"כ שעות אדם | תוצר מדיד |
| 1 | דוח מכין | 24/10/2021 | 50 | דוח מכין |
| 2 | חקירה ראשונית +לימוד עצמי של קורסים באינטרנט הקשורים לפרויקט +הכנת DATABASE המתאים לפרויקט | 1/1/2022 | 200 | לימוד פיתון +אלגוריתמי לימוד מכונה קשורים לפרויקט |
| 3 | דוח התקדמות | 16/1/2022 | 60 | דוח התקדמות של 25 עמוד לפחות |
| 4 | לימוד אלגוריתם SVM  לימוד אלגוריתם KNN  לימוד אלגוריתם NB  כתיבת התוכנה המתאימה + והרצת לימוד הנתונים בתוכנה | 25/6/2022 | 180 | בניית תוכנית והרצת נתונם וניתוח תוצאות |
| 5 | יום פרויקטים + הדגמה מעשית | 12/7/2022 | 30 | פוסטר + מצגת + POC |
| 6 | הכנת ספר הפרויקט  הכנה והגשה של מסמך RED | 1/9/2022 | 60 |  |
| 7 | הגנות | לימודי יום 12/7/2022  לימודי ערב 15/9/2022 | 8 | ספר פרויקט + פרויקט עובד |

**588 שעות**

תוצר מדיד:

תוצר הוא מה שהסטודנט בוחר להציג - מה שנבחר כתוצר של אבן הדרך. התוצר הוא משהוא מדיד, שלפני הפרויקט לא היה קיים והסטודנט ייצר אותו תוך כדי עבודתו על הפרויקט. על הסטודנט להתמקד בתיאור היצירה החדשה שנוצרה. מומלץ לציין את החידוש, ביחס למה שהיה בפרויקט לפני אבן דרך זו, וכיצד זה מקדם את הפרויקט.

**הערה – הצגת תוצר מדיד בטבלת אבני דרך במסגרת דוח המכין הוא מעיקרי הדוח.**

# נספח ג - טבלת משימות (לא חובה במסגרת דוח מכין - לפי החלטת המנחה)

במסגרת דוח ההתקדמות יש להכין טבלת משימות. (אם טבלת אבני הדרך היא רשימת פרקים, אזי טבלת משימות היא רשימת תתי פרקים). לכל אבן דרך, יש להציג את רשימת המשימות המובילות להשלמת אבן הדרך וקבלת התוצר המדיד (כולל הערכת משך הזמן הנדרש להשלמת כל משימה (טבלה המשימות היא טבלה נפרדת המכילה את כל המשימות יחדיו של כל אבני הדרך. מספור המשימות הוא בהתאם למספור אבני הדרך – 1.1, 1.2 .1.3 וכ"ו).

להלן רשימת העמודות שיש לכלול בטבלת המשימות:

• מספר משימה

• שם משימה

• מועד התחלה משוערך

• מועד סיום משוערך

• הקצעת שעות עבודה

• תוצר ביניים

• מועד סיום בפועל

• סה"כ שעות בפועל

# נספח ד

במידת הצורך, יש לצרף חתימה על מסמכים רלוונטיים של המרכז האקדמי רופין. הנחיות פרטניות תימסרנה בעתיד.

# נספח ה - גאנט (אופציונלי בהתאם לשיקול דעתו של המנחה)

Timeline

Description automatically generated