RECIBIDO EL 9 DE ABRIL DE 2021 - ACEPTADO EL 10 DE JULIO DE 2021

AN APPROXIMATION TO THE RISK FACTORS LINKED TO ALZHEIMER'S DISEASE

UNA APROXIMACIÓN A LOS FACTORES DE RIESGO RELACIONADOS CON LA ENFERMEDAD DE ALZHEIMER

Medina Cañizales Johan Sebastián[1]
Facultad de Ingeniería. Universidad Nacional de Colombia. Bogotá, Colombia.

Salcedo Parra, Octavio José[2]
Profesor Titular. Facultad de Ingeniería. Universidad Distrital Francisco José de Caldas. Bogotá, Colombia. Profesor de Planta, Departamento de Ingeniería de Sistemas e Industrial, Universidad Nacional de Colombia, Sede Bogotá

Rodríguez Miranda, Juan Pablo[3]
Profesor Titular. Facultad del Medio Ambiente y Recursos Naturales. Universidad Distrital Francisco José de Caldas. Bogotá, Colombia.

ABSTRACT

The objective of this project is to develop an application that integrates AI, machine learning and cloud computing technologies to reduce the user's chances of suffering from Alzheimer's disease in the future. The expectations regarding the application include that it can collect data about the most common risk factors in Alzheimer patients such as education level, hormonal and nutritional levels and lifestyle habits (smoking, sedentary lifestyle, alcohol, etc.) to prevent Alzheimer. It is also expected that it can determine the level of danger and the main risk factors that the user should be aware of. This software will keep running in cloud computing to ensure that it is always synchronized with the databases containing the factors described above, such as MEDLINE, PubMed, AlzForum (ALZRISK) and

1 Facultad de Ingeniería. Universidad Nacional de Colombia. Bogotá, Colombia. Correo electrónico: josmedinaca@unal.edu.co
ORCID: https://orcid.org/0000-0003-0867-4571
2 Profesor Titular. Facultad de Ingeniería. Universidad Distrital Francisco José de Caldas. Bogotá, Colombia. Profesor de Planta, Departamento de Ingeniería de Sistemas e Industrial, Universidad Nacional de Colombia, Sede Bogotá.Correo electrónico: osalcedo@udistrital.edu.co; ojsalcedop@unal.edu.co
ORCID: https://orcid.org/0000-0002-0767-8522
3 Profesor Titular. Facultad del Medio Ambiente y Recursos Naturales. Universidad Distrital Francisco José de Caldas. Bogotá, Colombia. Correo electrónico: jprodriguezm@udistrital.edu.co
ORCID: https://orcid.org/0000-0002-3761-8221

ADNI. The latter collects different data sources such as AIBL and DoD-ADNI.

It was found that the main risk factors that increase the probability of suffering the disease are: being overweight with a maximum value of 0.5 p-value when the diet followed is not balanced. A value of 0.3 p-value indicates that the diet is not followed. A deficit in physical activity is related to a p-value of 0.11. The years of formal studies also contribute by a significant value of 0.008 when said years are between 0 and 5. Stress from work can contribute with a maximum value of 0.006.

KEY-WORDS:

p-value, Alzheimer, risk factors, scrum, artificial intelligent, machine learning, dataset

RESUMEN

El objetivo de este artículo es desarrollar una aplicación que integre tecnologías de inteligencia artificial, aprendizaje automático y computación en la nube para reducir las posibilidades del usuario de padecer la enfermedad de Alzheimer en el futuro. Entre las expectativas con respecto a la aplicación se encuentra que pueda recopilar datos sobre los factores de riesgo más comunes en los pacientes de Alzheimer como el nivel educativo, los niveles hormonales y nutricionales y los hábitos de vida (tabaquismo, sedentarismo, alcohol, etc.) para prevenir el Alzheimer. También se espera que pueda determinar el nivel de peligro y los principales factores de riesgo que el usuario debe conocer. Este software seguirá ejecutándose en la computación en la nube para garantizar que siempre esté sincronizado con las bases de datos que contienen los factores descritos anteriormente, como MEDLINE, PubMed, AlzForum (ALZRISK) y ADNI. Este último recopila diferentes fuentes de datos como AIBL y DoD-ADNI.

PALABRAS CLAVE:

Alzheimer, factores de riesgo, scrum, inteligencia artificial, aprendizaje automático, conjunto de datos

I. INTRODUCTION

Alzheimer's disease is the most common form of dementia. It is estimated that every 66 seconds someone in the United States develops said disease and it is forecast that by year 2020 close to 14 million Americans could suffer from it [1].

Alzheimer begins slowly, affecting thought, memory and language. Over time, the symptoms worsen so it is urgent that people become aware of this disease and incorporate activities to their lifestyle for preventive reasons.

There are many factors that lead to Alzheimer's, from genetic factors (hereditary) [6] up to low nutritional levels [7] and even anxiety levels or sleep deprivation [9].

The following projects will be used as a reference:

• Detecting Alzheimer's disease on a Small Dataset: A Knowledge Transfer Perspective [10] uses a small population of people and tries to detect people with Alzheimer's through an algorithm.

• Deep Learning Framework for Alzheimer's Disease Diagnosis [11], which uses neural networks and Deep Learning algorithms to diagnose the disease through magnetic resonance imagery (MRI).

II. RELATED WORK

A. Factors associated with quality of life in patients with Alzheimer's disease. [2]

The purpose of this project is to find factors that affect the life of Alzheimer patients. It was concluded that several combinations of drugs led to depression, anxiety and troubles sleeping.

No information will be taken from this work.

B. The role of sleep deprivation and circadian rhythm disruption as risk factors of Alzheimer's disease [3]

This project looks for evidence that suggests that sleep deprivation and irregular cardiac rhythms increase the risk of suffering Alzheimer. It also delivers new data on the matter.

No information will be taken from this work.

C. Increased Alzheimer's risk during the menopause transition [4]

Two thirds of people who suffer from Alzheimer are women. This work studies the relationship between the different stages of menopause and the risk of developing the disease.

No information will be taken from this work.

D. Alzheimer's disease: nutritional status and cognitive aspects associated with disease severity [5]

This project studies various patients with Alzheimer and gathers the data concluding that, as the disease progresses, patients will exhibit a malnutrition and a more severe cognitive decay.

No information will be taken from this work.

E. Prevalence of variants of the apolipoprotein gen E (APOE) in adults in the urban area of Medellin (Antioquia) [6].

The purpose of this project consisted on finding the frequencies of the gen causing the disease in the population of Medellin, by considering data such as age, school years, gender and socioeconomic strata.

Demographic characteristics and clinical background of the population from low and middle-level income of the city of Medellin are taken.

F. Factors linked to cognitive decay in population under 65 years old: A systematic review [7]

The Purpose of this article is to determine the factors linked to the development of diseases such as age, socioeconomic strata, overweight, diabetes, educational level, among others.

The risk of bias in individual studies as well as evidence of social and economic factors in people under 65 years are taken.

G. Cholesterol, APOE genotype, and Alzheimer disease [8]

In this project, the relation of developing the disease with the level of cholesterol and other lipids and the APOE genotype are examined.

Different conclusions can be found in the article discussion.

H. Midlife Work-Related Stress Increases Dementia Risk in Later Life: The CAIDE 30-Year Study [9]

This project researches the relation between work-induced stress in a middle-aged person and the long-term development of dementia and Alzheimer.

The results of the analysis of the population, their sociodemographic characteristics, the relationship between work-induced stress and the subsequent development of dementia are taken.

III. METHODOLOGY

This project is developed in the agile SCRUM software tool. It is initially proposed to develop the software within 6 cycles (Sprint):

1. Research: The data to be used must be validated and the different sources of databases must be selected. Articles must be added or suppressed to determine the basic requirements of the app.

2. Planning: Different schemes to satisfy the requirements, the databases, languages to be used, platforms and the type of end user of the application.

3. Design: The screen browsed by the user are defined.

4. Production: The software prototype is created and released for testing.

5. Feedback: After testing, the failures are detected and corrected.

6. Deployment: The final version of the software is delivered with the corrected failures.

The preferred databases to use are:

• **AlzForum**

A website that gathers all the information related with Alzheimer, papers, diagnoses and treatment. Additionally, there are different databases that can be accessed freely.

AlzForum is comprised of the following databases:

-*AlzBiomarker*, captures two decades of research data on biomarkers linked to the disease.

-*AlzGene*, is a collection of studies on the genes causing the disease.

-*AlzPedia*: Is a quick reference tool that provides different terms related with the disease.

-*AlzRisk*: This database includes the studies that consider different risk factors (lifestyle habits and protein deficit) related with Alzheimer patients.

• **ADNI – Alzheimer's Disease Neuroimaging Initiative**

It is a collection of studies related with the disease whose purpose is to detect Alzheimer during its early stages and the follow-up of the disease in patients.

- **Dataset**

A dataset prototype with 50 rows and 9 columns is created that contains the data of 50 people including age, years of education, socioeconomic strata, work-related stress, smoking in a scale from 1 to 5, drinking alcohol in three levels (does not drink = 1, moderated drinker = 2, extreme drinker = 3). Physical activity measured in miles / day and whether a Mediterranean diet is followed.

```
age,yearseducation,wealth,stress,smoker,alcoholdrinker,weight,activity,diet
19,5,2,2,1,1,20,0.3,0
19,4,5,2,0,3,18,0.8,1
19,5,2,1,0,2,22,0.9,0
20,6,3,5,1,3,23,0.5,0
20,6,4,4,1,2,24,0.8,0
22,8,5,3,0,2,25,1.5,1
23,7,6,3,1,3,33,1.4,1
24,5,2,4,1,2,32,1.6,0
25,3,4,2,0,3,18,1.6,1
26,6,2,3,0,1,24,2,1
23,2,1,1,1,2,25,3,1
28,1,1,5,1,2,22,2,0
27,2,3,4,0,3,21,0.4,0
29,5,4,3,1,2,30,0.6,0
33,6,2,5,1,3,21,1.5,0
55,4,1,4,0,2,26,1.9,1
56,7,4,2,1,1,23,0.6,0
```

Source: Authors

ALGORITHMS

The prediction algorithm is adapted using artificial intelligence and machine learning through Python libraries. The algorithm is specialized in predicting values with some correlation with the current dataset. It is taken from DataQuest [12].

SELECTION OF THREE MAXIMUM VALUES IN AN ARRAY.

Furthermore, the *'Find the largest three elements in an array'* algorithm is integrated in order to find the three largest values in an array taken from GeeksforGeeks [13].

IV. DEVELOPMENT

This implementation reads the data within a previously built dataset and receives user information as inputs by keyboard. If the data indicated by the user are found in the program predictions, the user is alerted with a given risk percentage of suffering the disease.


```
# Import the pandas library.
import pandas
# Read in the data.
data = pandas.read_csv("data.csv")
# Print the names of the columns in games.
print(data.columns)

Index(['age', 'yearseducation', 'wealth', 'stress', 'smoker', 'alcoholdrinker',
       'weight', 'activity', 'diet'],
      dtype='object')
```

Source: Authors

- The data.csv file is loaded that contains the dataset to be used.

```
[ ] data.corr()["age"]

age                1.000000
yearseducation     0.305874
wealth            -0.060508
stress            -0.033201
smoker            -0.060618
alcoholdrinker     0.026972
weight             0.236100
activity           0.050061
diet              -0.042665
Name: age, dtype: float64
```

Source: Authors

- Data correlation is measured.

```
[ ] data.corr()["yearseducation"]

age                0.305874
yearseducation     1.000000
wealth             0.323693
stress            -0.147995
smoker            -0.014724
alcoholdrinker    -0.052030
weight            -0.085968
activity          -0.060770
diet              -0.039345
Name: yearseducation, dtype: float64
```

Source: Authors

```
[ ] data.corr()["activity"]

age                0.050061
yearseducation    -0.060770
wealth            -0.175913
stress            -0.031454
smoker             0.040469
alcoholdrinker    -0.105174
weight            -0.041916
activity           1.000000
diet               0.430240
Name: activity, dtype: float64
```

```
[ ] data.corr()["diet"]

age               -0.042665
yearseducation    -0.039345
wealth            -0.122358
stress             0.050865
smoker             0.019592
alcoholdrinker     0.153100
weight             0.028018
activity           0.430240
diet               1.000000
Name: diet, dtype: float64
```

Source: Authors

```
[ ]  # Get all the columns from the dataframe.
     columns = data.columns.tolist()

     # Store the variable we'll be predicting on.
     target = "age"
     target2 = "yearseducation"
     target3 = "wealth"
     target4 = "stress"
     target5 = "smoker"
     target6 = "alcoholdrinker"
     target7 = "weight"
     target8 = "activity"
     target9 = "diet"
```

Source: Authors

- The column titles of the dataset are loaded and different variables are allocated to be subsequently used in the predictions of their respective values.

```
# Import a convenience function to split the sets.
from sklearn import model_selection

# Generate the training set.  Set random_state to be able to replicate results
train = data.sample(frac=0.8, random_state=1)
# Select anything not in the training set and put it in the testing set.
test = data.loc[~data.index.isin(train.index)]
# Print the shapes of both sets.
print(train.shape)
print(test.shape)
```


Source: Authors

- The datasets are created to train the program so it can predict values.

```
# Import the linearregression model.
from sklearn.linear_model import LinearRegressic

# Initialize the model class.
model = LinearRegression()
model2 = LinearRegression()
model3 = LinearRegression()
model4 = LinearRegression()
model5 = LinearRegression()
model6 = LinearRegression()
model7 = LinearRegression()
model8 = LinearRegression()
model9 = LinearRegression()

# Fit the model to the training data.
model.fit(train[columns], train[target])
model2.fit(train[columns], train[target2])
model3.fit(train[columns], train[target3])
model4.fit(train[columns], train[target4])
model5.fit(train[columns], train[target5])
model6.fit(train[columns], train[target6])
model7.fit(train[columns], train[target7])
model8.fit(train[columns], train[target8])
model9.fit(train[columns], train[target9])
```

Source: Authors

- 9 models are generated, one for each column. A linear regression is carried out for each model and the information is organized so it matches the data used for training.

```
# Import the scikit-learn function to compute error.
from sklearn.metrics import mean_squared_error
# Generate our predictions for the test set.
predictions = model.predict(test[columns])
predictions2 = model2.predict(test[columns])
predictions3 = model3.predict(test[columns])
predictions4 = model4.predict(test[columns])
predictions5 = model5.predict(test[columns])
predictions6 = model6.predict(test[columns])
predictions7 = model7.predict(test[columns])
predictions8 = model8.predict(test[columns])
predictions9 = model9.predict(test[columns])
# Compute error between our test predictions and the actual values.
mean_squared_error(predictions, test[target])

6.437104986603456e-29
```

Source: Authors

- The data predictions are generated for each column as well as the error, which has an order of 10^{-29}, making it fairly low.

```
print("Predicciones para edad")
print(predictions)

Predicciones para edad
[19. 22. 25. 26. 28. 27. 55. 56. 46. 24.]

print("Predicciones para años de educacion")
print(predictions2)

Predicciones para años de educacion
[ 5.  8.  3.  6.  1.  2.  4.  7.  9. 16.]

print("Predicciones para estrato ")
print(predictions3)

[2. 5. 4. 2. 1. 3. 1. 4. 1. 1.]

print("Predicciones para nivel de estres-relacionado con el trabajo ")
print(predictions4)

Predicciones para nivel de estres-relacionado con el trabajo
[2. 3. 2. 3. 5. 4. 4. 2. 4. 3.]
```

Source: Authors

- The values are predicted for each risk factor.

```
name = input("¿Cómo se llama? ")
yearsEducation = float(input("¿Cuantos años ha estudiado? "))
wealth = float(input("Estrato socioeconomico de 1-6 :"))
stress = float(input("Nivel de estres en su trabajo de 1-5 : "))
smoker = float(input("¿Es fumador? Si(1)/No(0) : "))
alcoholdrinker = float(input("Nivel de consumo de alcohol 1(nulo)/2(Casual)/3
weight = float(input ("Su indice de peso segun BMI: "))
activity = float(input ("Nivel de actividad fisica diario de 0.0 a 2.0: "))
diet = float(input("¿Sigue una dieta mediterranea? Si(1)/No(0) : "))
```

Source: Authors

- The user data is loaded by keyboard.

```
for x in predictions2:
  if(round(x) == yearsEducation and yearsEducation >=0

    pvalueCount += 0.008
    break
  elif((round(x) == yearsEducation and yearsEducation

      pvalueCount += 0.004
      break
  elif ((round(x) == yearsEducation and yearsEducation

        pvalueCount += 0.002
        break
valorEstd = pvalueCount
print ("pvalue para años de educacion:  ",pvalueCount
```

Source: Authors

- The p-values are computed for different risk factors and added in a cumulative manner.

```
def print3largest(arr, arr_size):

  # There should be atleast three
  # elements
  if (arr_size < 3):

    print(" Invalid Input ")
    return

  third = first = second = -sys.maxsize

  for i in range(0, arr_size):

    # If current element is greater
    # than first
    if (arr[i] > first):

      third = second
      second = first
      first = arr[i];aux1 = i

    # If arr[i] is in between first
    # and second then update second
    elif (arr[i] > second):

      third = second
      second = arr[i];aux2 = i

    elif (arr[i] > third):
      third = arr[i];aux3 = i
```


Source: Authors

- The highest p-values are chosen as well as the risk factors to inform the user.

```
¿Como se llama? Sombra
¿Cuantos años ha estudiado? 6
Estrato socioeconomico de 1-6 :4
Nivel de estres en su trabajo de 1-5 : 4
¿Es fumador? Si(1)/No(0) : 0
Nivel de consumo de alcohol 1(nulo)/2(Casual)/3(Extremo)3
Su indice de peso segun BMI: 22
Nivel de actividad fisica diario de 0.0 a 2.0: 1.6
¿Sigue una dieta mediterranea? Si(1)/No(0) : 0
```

Source: Authors

- The testing values are entered in the software.

```
pvalue para años de educacion:   0.004
pvalue para estrato:   0.0077
pvalue para stress-trabajo:   0.004
pvalue para variable de fumador:   0
pvalue para bebedores de alcohol:  0.066
pvalue para masa corporal (peso/BMI):  0.001
pvalue para actividad fisica:  0.03
pvalue para dieta-mediterranea:  0.063
Los factores de riesgo mas altos son  Alcohol: 0.066
 Dieta mediterranea: 0.063 Actividad fisica: 0.03
Porcentaje total de sufrir la enfermedad basado
en los factores de riesgo:  17.57
```

Source: Authors

- The results are printed and the user is informed of the percentage of suffering the disease and the highest-ranking risk factors.

V. Discussion of results

A. Prevalence of variants in the apolipoprotein E (APOE) in adults of the urban area of Medellin (Antioquia) [6].

Results: A total of 694 individuals were analyzed with a high rejection frequency of the disease in the socioeconomic strata 4 (middle), 5 (middle-high) and 6 (high).

Tabla 1 – Características demográficas y antecedentes clínicos de los adultos de población general de estratos bajos y medios-bajos de la ciudad de Medellín, 2010

Pacientes, n	964
Edad (años)	45,07 ± 16,17 (18-93)
Escolaridad (años)	8,41 ± 4,35 (0-23)
Mujeres	645 (66,9)
Estado civil	
Soltero	348 (36,1)
Casado	292 (30,3)
Unión libre	181 (18,7)
Separado o divorciado	74 (7,7)
Viudo	69 (7,2)
Estrato socioeconómico	
1 (muy bajo)	108 (11,2)
2 (bajo)	413 (42,8)
3 (medio bajo)	333 (34,5)
4 (medio)	74 (7,7)
5 (medio alto)	33 (3,4)
6 (alto)	3 (0,3)

Los valores expresan n (%) o media ± desviación estándar (intervalo).

Source: Prevalence of variants in the apolipoprotein gen E (APOE) in adults in urban Medellin (Antioquia) [6]

Tabla 3 – Evidencia de factores sociales, económicos o conductuales y genético en menores de 65 años

Factor	País	Muestra	Tiempo de seguimiento (años
Educación[22]	Países Bajos	1.823	6
Tabaquismo[8,15,17,27]	Estados Unidos, Reino Unido	10.963, 1.920, 3.004	6, 6, 20, 5
Estatus socioeconómico[30]	Reino Unido	1.744	5,5
Ocupación (h/semana)[23]	Reino Unido	2.214	5
Estado civil[24]	Suecia	1.882	5
Estimulación cognitiva dentro y fuera del trabajo[28]	Francia	3.123	10
Estado cognitivo en la adolescencia[30]	Reino Unido	2.058	10
Actividad física[15,27]	Países Bajos, Reino Unido	1.904, 3.004	5, 20
Dieta[27]	Reino Unido	3.004	20
Factores genéticos, APOE-ε4[illegible]	Estados Unidos	11.580, 180, 1.920	6, 2,75, 6

Source: Factors related with the cognitive decay in population under 65 years old: a systematic review [7]

The sociodemographic characteristics can be seen in **table 1**. The allele E3 had the highest frequency in the population (92%; IC95%, 88,6%-94,5%) and the alleles E2 and E4 had a similar prevalence between them.

CONCLUSION OF RESULTS

The results of the project [6] determined genetic, socioeconomic and demographic patterns on how the disease has developed in this region (Medellin) throughout time.

B. Factors linked to the cognitive decay in people under 65 years: a systematic review [7]

Results: The median of participants per study was 2.214. The median of the age average was 49.5 (35 – 60) years. The median of the percentage of women included was 51% (23% - 61%) and the median of school years was 33.4% (19.4 – 58%).

CONCLUSION OF THE RESULTS

The results of the project [7] aim to find factors linked to Alzheimer's such as diabetes, overweight, educational levels, physical activity, cognitive stimulation, quality of the diet in people under 65 years old. Nonetheless, the data is affected by a low quality of the evidence.

C. Midlife Work-Related Stress Increases Dementia Risk in Later Life: The CAIDE 30-Year Study [9]

Results: The following table shows the sociodemographic and clinical characteristics of the population analyzed in the study. The study shows that the allele APOE ε4 and other risk factors are present in the subjects. 1,511 patients were examined, 61 were diagnosed with dementia and 48 with Alzheimer.

Table 1. Sociodemographic and Clinical Characteristics of the Entire Population

Characteristics	p Value
Age at midlife	<.001
Age at late life*	N.S.
Sex	
Women	
Men	N.S.
Education (years)	<.001
APOE ε4 allele (nonparticipants n = 29)	
Noncarrier	
Carrier	N.S.
Work-related stress (range 1–5)	.006
Midlife BMI	<.001
Midlife systolic blood pressure (mmHg)	<.001
Midlife cholesterol	.002
Type of occupation	
White-collar	
Other	<0.001
Smoker	
No	
Yes	N.S.

Source: Midlife Work-Related Stress Increases Dementia Risk in Later Life: The CAIDE 30-Year Study [9]

The results of the project [9], aim to find the relationship between work-induced stress and the chances of developing Alzheimer in the future. The results show that people are more susceptible to suffer from the disease if they have less years of formal education, high levels of stress from work, high blood pressure and high levels of cholesterol. Most of the people had jobs that required significant physical effort (farmers, miners, workers in construction, housewives and others).

D. Approximation to risk factors linked to Alzheimer

The Project ‘Approximation to risk factors linked to Alzheimer’ proposes to determine a certain percentage of suffering the disease based on risk factors to reduce it over time and correct the risk factors in the eventual development of the disease. The results are the highest risk factors presented by the user in his lifestyle and an estimated percentage of suffering from the disease in the future. It is known that the cognitive decay starts 20 years after diagnosing the disease. It is intended that the user can identify and correct lifestyle habits that lead to high risk factors.

VI. CONCLUSIONS

The results of the project aim for the early detection of Alzheimer, identifying the highest risk factors of the user and delivering a percentage of suffering the disease, so that the user can correct his lifestyle. Nonetheless, the databases used for the development of the project were very limited so the results are an approximation of the real value. The most critical risk factors (the ones that contribute the most to the chances of suffering from the disease) are: years of formal study (scholarship level, **low** contributes to 0.008p-value, **medium** contributes to 0.004 p-value, **high** contributes to 0.002p-value), overweight (**BMI**<**20kg/m^2** contributes with 0.001p-value, **25kg/m^2** <**BMI**<**30kg/m^2** contributes with 0.44p-value, **BMI**>=**30kg/m^2** contributes with 0.5p-value), physical activity (**activity**<**0.25** contributes with 0.11p-value, **0.25<activity**<=**1** contributes with 0.09p-value, **activity**>**1** contributes with 0.03p-value) and balanced diet patterns. This project was based on the data of a **Mediterranean** diet (**Following the diet** contributes with 0.006p-value. **Not following diet**, contributes with 0.3p-value).

One of the biggest problems regarding the disease is the late diagnosis since the symptoms may show 20 years later, when the disease has significant developed. In that sense, it is concluded that the early diagnosis and/or correction of risk factors are crucial to reducing the possible development of the disease. This first approximation for the prevention of Alzheimer’s can be improved in future work.

VII. BIBLIGRAPHICS REFERENCES

[1] Why the number of Americans with Alzheimer's could more than double by 2050, 2019. Disponible online https://to.pbs.org/2W5bBje

[2] Factors associated with quality of life in patients with Alzheimer's disease, Coralie Barbe, Damien Jolly, Isabella Monrrone, 2018. https://doi.org/10.1186/s12877-018-0855-7

[3] The role of sleep deprivation and circadian rhythm disruption as risk factors of Alzheimer's disease, Hao Wu, Sophie Dunnett, Yuen-Shan Ho, 2019. https://doi.org/10.1016/j.yfrne.2019.100764

[4] Increased Alzheimer's risk during the menopause transition, Lisa Mosconi, Aneela Rahman, Ivan Diaz, Xian Wu, 2018. https://doi.org/10.1371/journal.pone.0207885

[5] Alzheimer's disease: nutritional status and cognitive aspects associated with disease severity, Lineau Correa, Gloria Souza, Julia Laura Delbue, Barbosa Nascimento, 2018. http://dx.doi.org/10.20960/nh.2067

[6] Prevalencia de variantes en el gen de la apolipoproteína E (APOE) en adultos de la población general del área urbana de Medellín (Antioquia), Arango Viana Juan Carlos, Palacio Carlos, García Valencia, 2014. http://www.redalyc.org/articulo.oa?id=80631556004

[7] Factores asociados con el declive cognitivo en población menor de 65 años. Una revisión, Lopera Restrepo Francisco, Aguirre Camilo, Giraldo Arango Diana, Jaimes Barragan Fabian, 2014. http://www.redalyc.org/articulo.oa?id=80631556008

[8] Cholesterol, APOE genotype, and Alzheimer disease, K. Hall, J. Murrell, R. Evans, 2006. https://doi.org/10.1212/01.wnl.0000194507.39504.17

[9] Midlife Work-Related Stress Increases Dementia Risk in Later Life: The CAIDE 30-Year Study, Shireen Sindi, Goran Hagman, Krister Hakansson, 2017. https://doi.org/10.1093/geronb/gbw043

[10] Detecting Alzheimer's Disease on Small Dataset: A Knowledge Transfer Perspective, Wei Li, Yifei Zhao, Xi Chen, 2019. http://bit.ly/2xc6SkT

[11] Deep Learning Framework for Alzheimer's Disease Diagnosis, Chiyu Feng, Ahmed Elazab, Peng Yang, 2019. http://bit.ly/2Lh7le2

[12] Learn Python Programming and Machine Learning, DataQuest, http://bit.ly/30sA5Vd

[13] Find the largest three elements in an array, GeeksForGeeks, http://bit.ly/2Jv2nsM