Calorie Burnt Prediction

Mini Project KCS 752/ 7th Sem

STUDENT NAME

Anshita Sharma

(2000320120035)

Outline

- Introduction
- Existing System
- Proposed System
- Objectives
- Tools and Technologies used
- Methodology Used
- Implementation
- Results
- Limitations and Future Enhancements
- Conclusions
- Reference

Introduction

- Usually, when people think of calories, they only think of food or weight loss. The measurement can be used to evaluate many energy releasing systems unrelated to the human body The amount of energy required by the body to perform a task is the number of calories considered from the point of view of the human body.
- Systems unrelated to the human body The amount of energy required by the body to perform a task is the number of calories considered from the point of view of the human body.
- The variables used here are the time scale a person exercises, average heart rate per minute, and temperature. Then add the person's height, weight, gender, and age to predict how much energy that person is burning. Parameters that can be taken into account are exercise time, average heart rate per minute, temperature, height, weight and gender. The XG Boost machine learning regression algorithm is used to predict calories burned based on exercise time, temperature, height, weight, and age.

Existing System

The existing systems for predicting calories burnt are based on a combination of factors, data, and algorithms. These systems aim to estimate the number of calories an individual burns during a specific activity or throughout the day. Here are some components and methods commonly used in existing systems for predicting calories burnt:

- **Biometric Data**: These systems often start with the user inputting biometric data, such as age, gender, weight, height, and, in some cases, body composition measurements. These factors play a crucial role in determining an individual's basal metabolic rate (BMR) - the calories burned at rest.
- **Activity Data**: Users typically provide information about their physical activities. This includes the type of exercise, duration, and intensity. In some advanced systems, data from wearable fitness devices (like fitness trackers) are used to monitor and record real-time activities.
- **Heart Rate Monitoring**: Some systems employ heart rate data to estimate calorie expenditure. Heart rate can be a more accurate indicator of energy expenditure during exercise because it reflects the effort exerted.

Proposed System

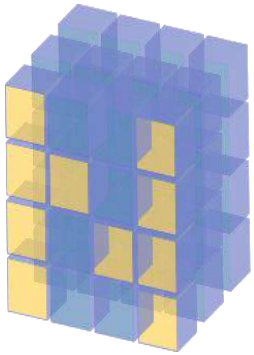
- The proposed system in the study is a machine learning system that uses attributes as input to predict the approximate calories burnt after exercise. It aims to motivate people to exercise more and track their daily growth.
- The system is trained on a dataset of more than 15,000 data points and has a Mean Absolute Error (MAE) of 1.48, which is expected to improve over time for better results.
- Factors affecting calories burnt during exercise include age, where as age increases, the need for calories decreases due to less physical activity and changes in metabolism.
- Other factors that affect calories burnt include body temperature, as higher body temperature leads to more energy being used to cool down, resulting in calorie burn.
- The duration of exercise also directly affects the calories burned, as longer durations lead to increased body temperature and extra calorie burn.

Objectives

- The objective of the study is to develop a machine learning system that can predict the calories burnt during exercise based on input attributes. The system aims to motivate individuals to exercise more by providing them with an approximate measure of their calorie burn and showing their daily growth.
- The goal is to provide a quick and convenient solution for individuals who want to check their level of improvement and track their calorie burn after exercise.
- The system is designed to take into account various factors that affect calories burnt, such as age, body temperature, duration of exercise, and height.
- By accurately predicting calories burnt, the system can help individuals make informed decisions about their exercise routines and encourage them to engage in more physical activity.

Tools and Technologies Used

- **Python**
- **NumPy**
- **Pandas**
- **Matplotlib**
- **Seaborn**
- **Scikit-learn**



NumPy

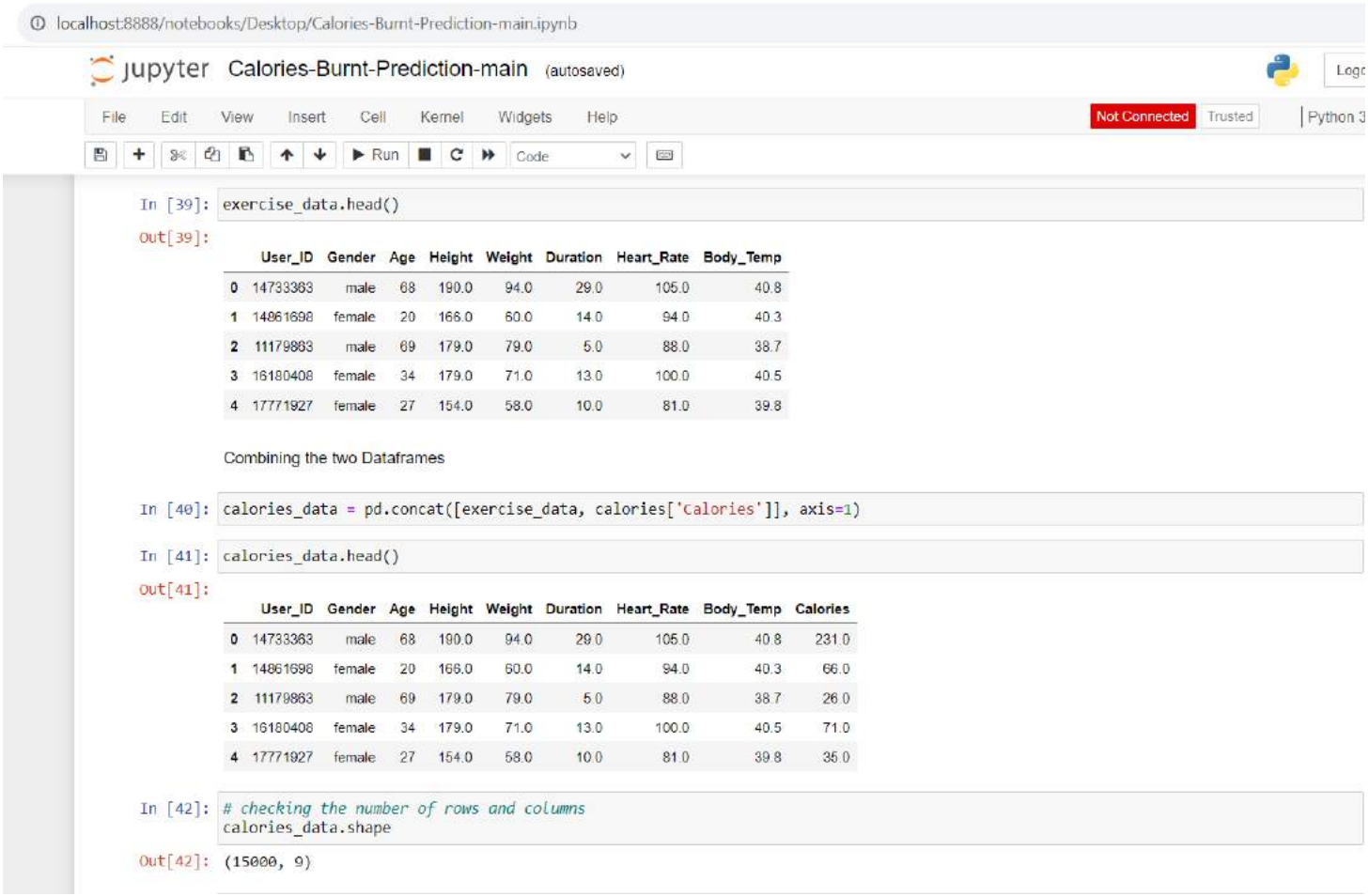
Methodology Used

- **Challenges Identification**
- **Data Loading and Understanding**
- **Data Description**
- **Imbalance Analysis**
- **Data Visualization**
- **Data Splitting**
- **Model Evaluation**
- **Confusion Matrix Visualization**
- **Comparison with Other Algorithms**

Methodology Used

- **Data collection :**Data retrieval is the first step. Kaggle is the data sto
  It is loaded into the Collab program. The information collected is bot
  categorical and numerical. Evaluated the model's performance using various metrics such as accuracy, precision, recall, F1-score, and Matthews correlation coefficient.
- **Data Preprocessing :**15,000 instances and 7 data attributes are contained in two csv files ("exercise.csv" and "calorie.csv." Each person's attributes are included in the Kaggle data collection. Including their size , weight, gender, age, exercise duration, heart rate, and body temperature.
- **Data Analysis:** Colab, the platform used for the processing, requires the upload of two dataset csv files ("exercise.csv" and "calorie.csv"). The average body temperature is 40. People who exercise will have a higher body temperature. Heart rate and coronary temperature were the most important results of this analysis. The data is then visualized using a few tables and charts.
- **Evaluation** This dataset was analyzed to make predictions about how many calories were burned based on exercise duration as well as factors such as age, gender, body temperature and heart rate at different time points. different points during exercise. We are looking for a machine learning model with lower mean absolute error that produces more accurate results using these machine learning methods.

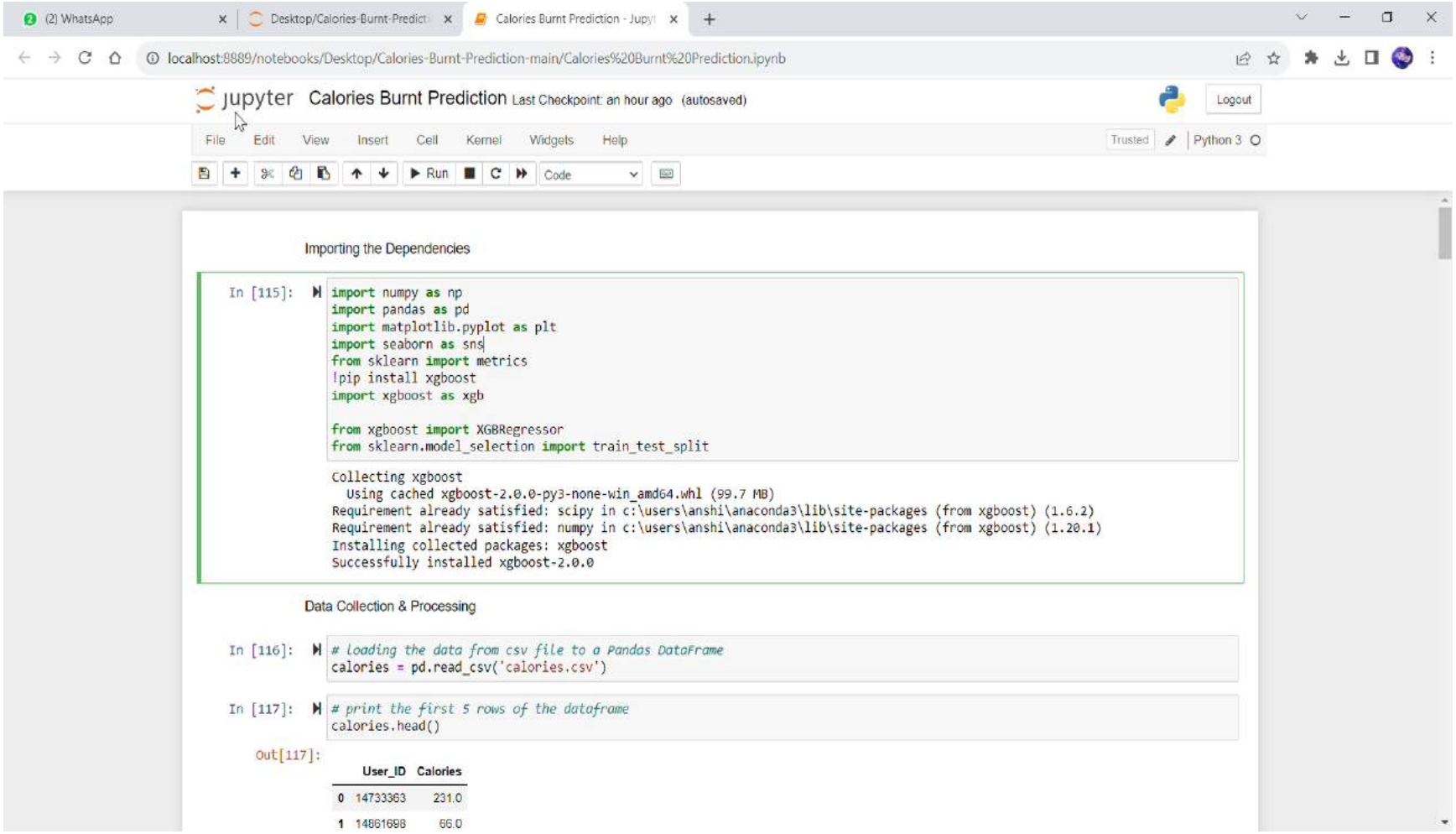
Implementation



	User_ID	Gender	Age	Height	Weight	Duration	Heart_Rate	Body_Temp
0	14733363	male	68	190.0	94.0	29.0	105.0	40.8
1	14861698	female	20	166.0	60.0	14.0	94.0	40.3
2	11179863	male	69	179.0	79.0	5.0	88.0	38.7
3	16180408	female	34	179.0	71.0	13.0	100.0	40.5
4	17771927	female	27	154.0	58.0	10.0	81.0	39.8

	User_ID	Gender	Age	Height	Weight	Duration	Heart_Rate	Body_Temp	Calories
0	14733363	male	68	190.0	94.0	29.0	105.0	40.8	231.0
1	14861698	female	20	166.0	60.0	14.0	94.0	40.3	66.0
2	11179863	male	69	179.0	79.0	5.0	88.0	38.7	26.0
3	16180408	female	34	179.0	71.0	13.0	100.0	40.5	71.0
4	17771927	female	27	154.0	58.0	10.0	81.0	39.8	35.0

Implementation



Implementation

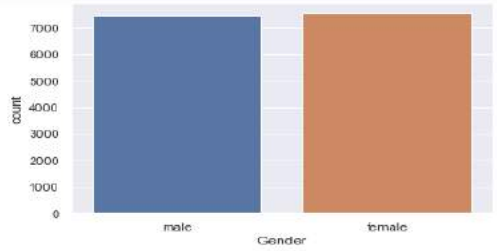
ost:8888/notebooks/Desktop/Calories-Burnt-Prediction-main.ipynb

jupyter Calories-Burnt-Prediction-main (autosaved)

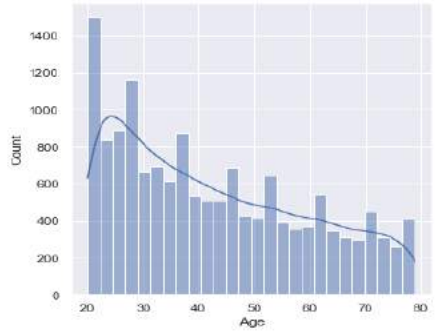
File Edit View Insert Cell Kernel Widgets Help

Not Connected Trusted Python 3

Run Code

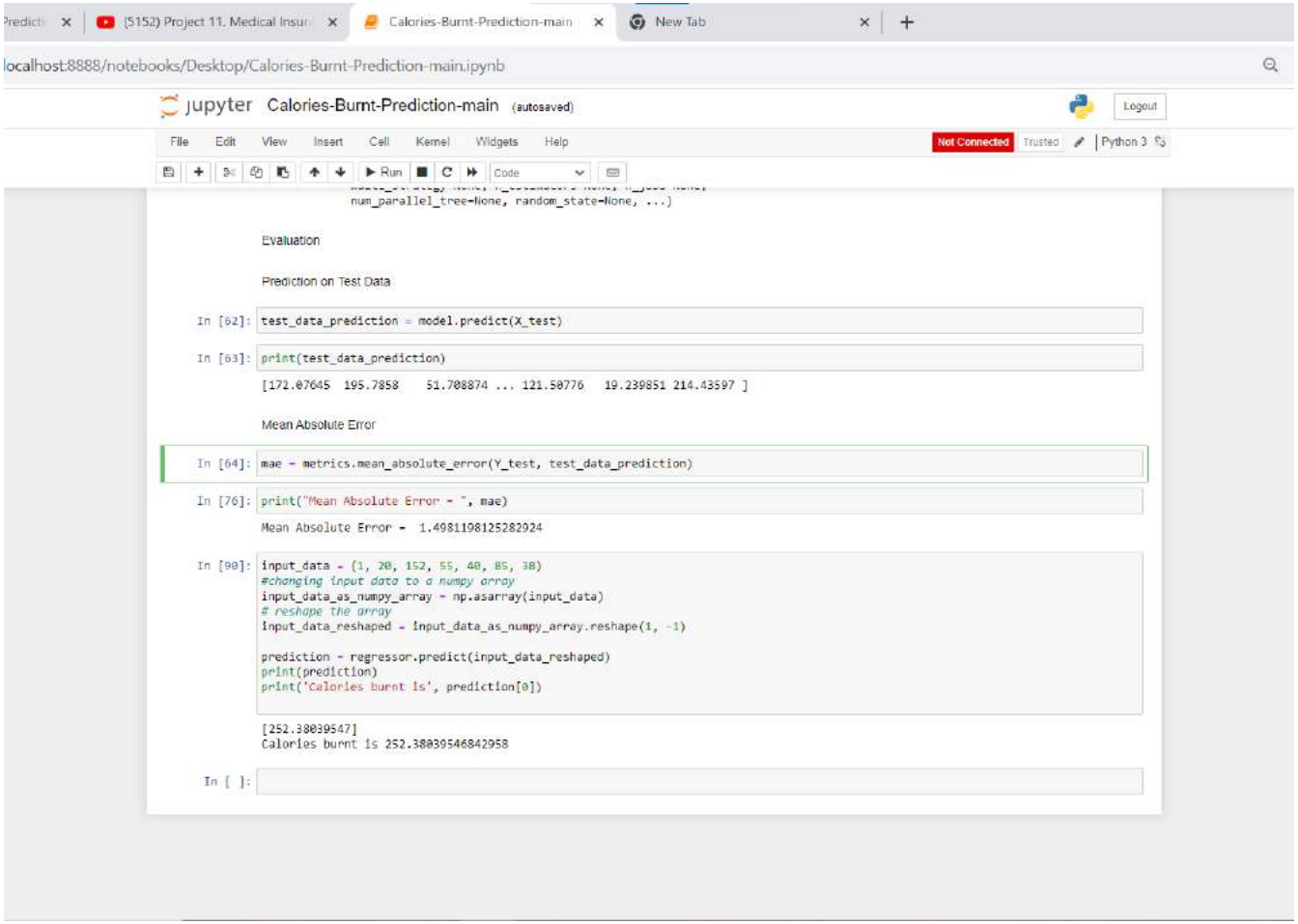


In [48]:
```
# finding the distribution of "Age" column
sns.displot(data=calories_data, x="Age", kde=True)
plt.show()
```



In [49]:
```
# finding the distribution of "Height" column
sns.displot(data=calories_data, x="Height", kde=True)
plt.show()
```

Result


(5152) Project 11. Medical Insur
Calories-Burnt-Prediction-main
New Tab
localhost:8888/notebooks/Desktop/Calories-Burnt-Prediction-main.ipynb
jupyter Calories-Burnt-Prediction-main (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Connected
Trusted
Python 3
Run
Code
num_parallel_tree=None, random_state=None, ...)
Evaluation
Prediction on Test Data
In [62]: test_data_prediction = model.predict(X_test)
In [63]: print(test_data_prediction)
[172.07645 195.7858 51.708874 ... 121.50776 19.239851 214.43597]
Mean Absolute Error
In [64]: mae = metrics.mean_absolute_error(Y_test, test_data_prediction)
In [76]: print("Mean Absolute Error = ", mae)
Mean Absolute Error = 1.4981198125282924
In [90]: input_data = (1, 20, 152, 55, 40, 85, 38)
#changing input data to a numpy array
input_data_as_numpy_array = np.asarray(input_data)
# reshape the array
input_data_reshaped = input_data_as_numpy_array.reshape(1, -1)
prediction = regressor.predict(input_data_reshaped)
print(prediction)
print('Calories burnt is', prediction[0])
[252.38039547]
Calories burnt is 252.38039546842958
In [ ]:

Limitations

- ✓ **Limitations of the Study:**
- ❖ The study only focuses on predicting burned calories using supervised machine learning regression algorithms, without considering other factors that may affect calorie burn, such as individual metabolism or exercise intensity.
- ❖ The accuracy of the regression models was evaluated based on the result of model testing after ten iterations, but it is unclear if this is sufficient to determine the overall performance and generalizability of the models.

Future Enhancements

- Future research could consider incorporating additional variables, such as heart rate or exercise duration, to improve the accuracy of calorie burn prediction models.
- The study could also explore the use of other machine learning algorithms or ensemble methods to further enhance the prediction accuracy.
- It would be beneficial to validate the performance of the regression models on a larger and more diverse dataset to ensure their applicability to a wider population.

Conclusions

- XGB Regressor produces results more accurate results. The mean absolute error indicates that the absolute error should be as small as possible. It is nothing more than the difference between the observed values and the values predicted by the models. 1.49 is a good value for the mean absolute value that XGB Regressor gives us. The error rate is quite low. Therefore, we can say that XG Boost Regressor is the best model to predict calories consumed. The flexibility of the proposed technique can also be improved with variations. In this study, we focused on seven main factors that influence how many calories our bodies burn, but there are other factors that play a role as well. It is also important to understand how many calories we consume if we want to stay healthy and fit. Alternatively, ML can be used to build this (machine learning). A user interface is also required for the user to enter their values and get results showing how many calories they have burned. Moreover, we can create a fully functional app with all these features and our recommended diet and exercise program.

Reference

- [1] Smola, A., & Vishwanathan, S. V. N. (2008). Introduction to machine learning. Cambridge University.
- [2] MacKay, D.J, & Mac Kay, D. J. (2003). Information theory, inference and learning algorithms.Cambridge University press.
- [3] Mitchell, T. M. (1999). Machine learning and data mining. Communications of the ACM
- [4] https://www.kaggle.com/fmendes/fmendesdat263xdemos.
- [5] https://machinelearningmastery.com/xgboost-for- regression/
- [6] https://www.medicalnewstoday.com/articles/319731#factors-influencing-daily-calorie-burn-and-weight-loss
- [7] https://zenodo.org/record/6365018
- [8] https://devhadvani.github.io/calorie.htm
- [9] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5496172
- [10] https://www.jetpac.com/

THANK YOU