Introduction:

The current state of AI is quite advanced, and its potential applications across different industries are numerous. Here are some examples of the potential applications of AI in different industries:

1. Healthcare: AI can be used to analyze large amounts of medical data to identify patterns and make more accurate diagnoses. It can also be used to develop personalized treatment plans and predict the effectiveness of different treatments.
2. Finance: AI can be used to analyze financial data, detect fraud, and make predictions about market trends. It can also be used to develop personalized investment portfolios and improve risk management strategies.
3. Manufacturing: AI can be used to optimize production processes, reduce waste, and improve product quality. It can also be used to develop predictive maintenance schedules and reduce downtime.
4. Retail: AI can be used to analyze customer data, personalize recommendations, and improve supply chain management. It can also be used to optimize pricing strategies and improve customer service.
5. Transportation: AI can be used to optimize routes, reduce congestion, and improve safety. It can also be used to develop autonomous vehicles and improve logistics and supply chain management.
6. Education: AI can be used to personalize learning experiences, develop adaptive learning platforms, and identify areas where students need additional support.

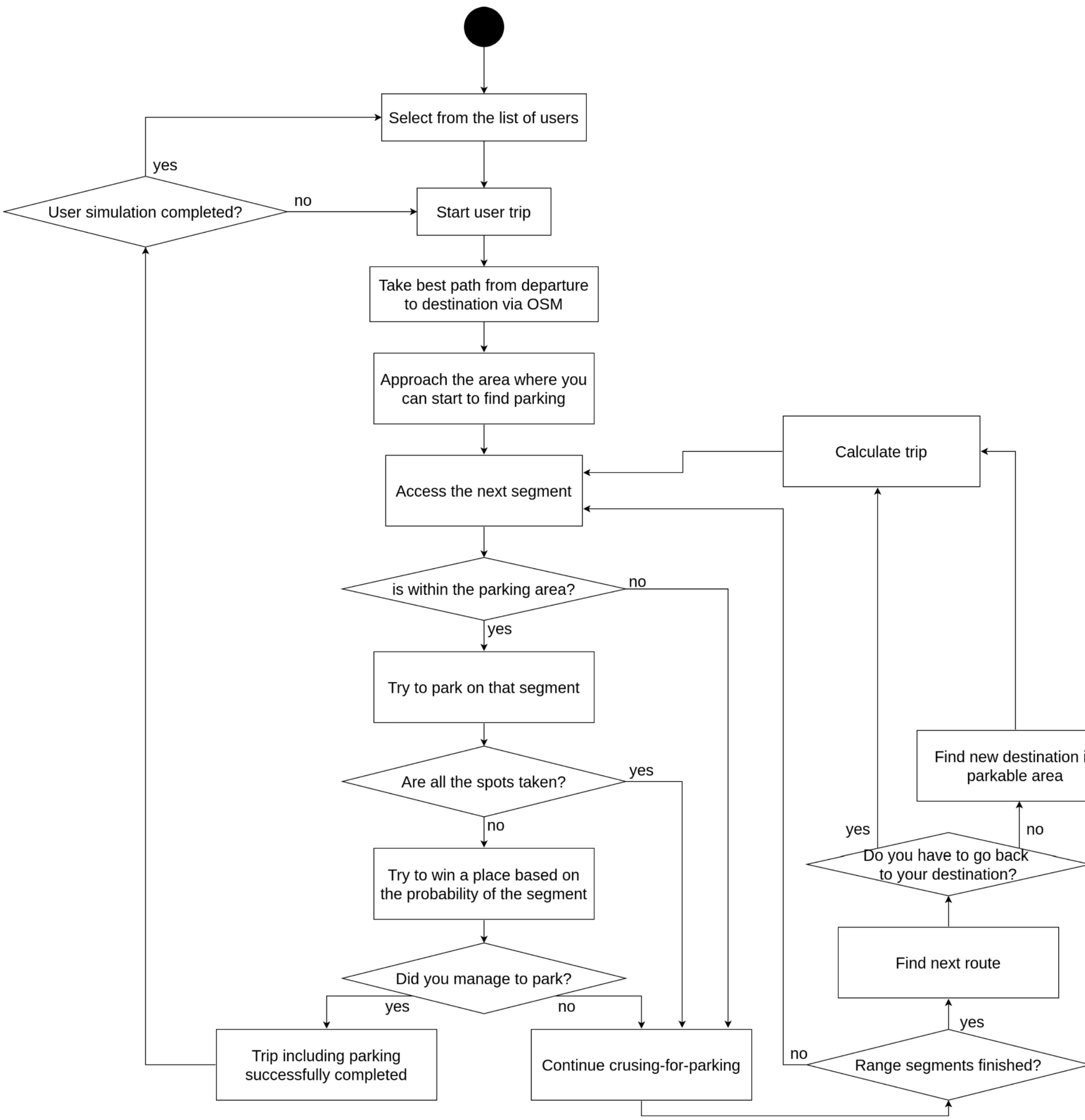
Background :

Overall, the potential applications of AI are vast and varied, and it has the potential to transform many industries in the coming years. However, it is important to ensure that AI is developed and deployed in an ethical and responsible manner to ensure that it benefits society as a whole.
Cruising-for-parking in an urban area is a time-consuming and frustrating activity. We present four machine learning-based models to predict the parking availability of street segments in an urban area on a three-level scale, which navigator and smart-parking apps can exploit to ease and reduce the cruising phase. The models were trained with data generated by a cruising-for-parking simulator that we developed, replicating four parking behavior types (workers, residents, buyers, and visitors). The generated data is comparable to that collectible with smartphones' sensors. We simulated 40 users moving for 200 weeks in the city area of San Giovanni in Rome. We collected information about users' parking, unparking, and cruising actions over considered road segments at different time slots. Once a significant amount of trips were collected, we extracted ten features for each road segment at a given time slot. With the obtained dataset, which contained 761 samples, we trained and compared four supervised machine learning models that receive the history of a segment and, in return, classify the Parking Availability Level of the segment as Green, Yellow or Red. The four models were further evaluated in a different city area, San Lorenzo, and obtained very accurate results. We can

predict parking availability with an accuracy above 97% for all the street segments where we collected 30 or more user actions, confirming the robustness of the simulator in generating synthetic cruising-for-parking data and the suitability of designing a Parking Availability Classifier (PAC) based on data collectible by smartphones.

**Keywords:** **parking availability**; **machine learning**; **artificial neural network**; **user-centered artificial intelligence**; **HCI**

Methodology:




The results achieved by the models in this test case were very high, reaching **97%** of accuracy. From the confusion matrices (**Figure 11**, **Figure 12**, **Figure 13** and **Figure 14**), it is also possible to note that errors, if any, always occur between yellow and another color, which is a less severe problem. It never happens that a green label is confused with a red one or vice versa.

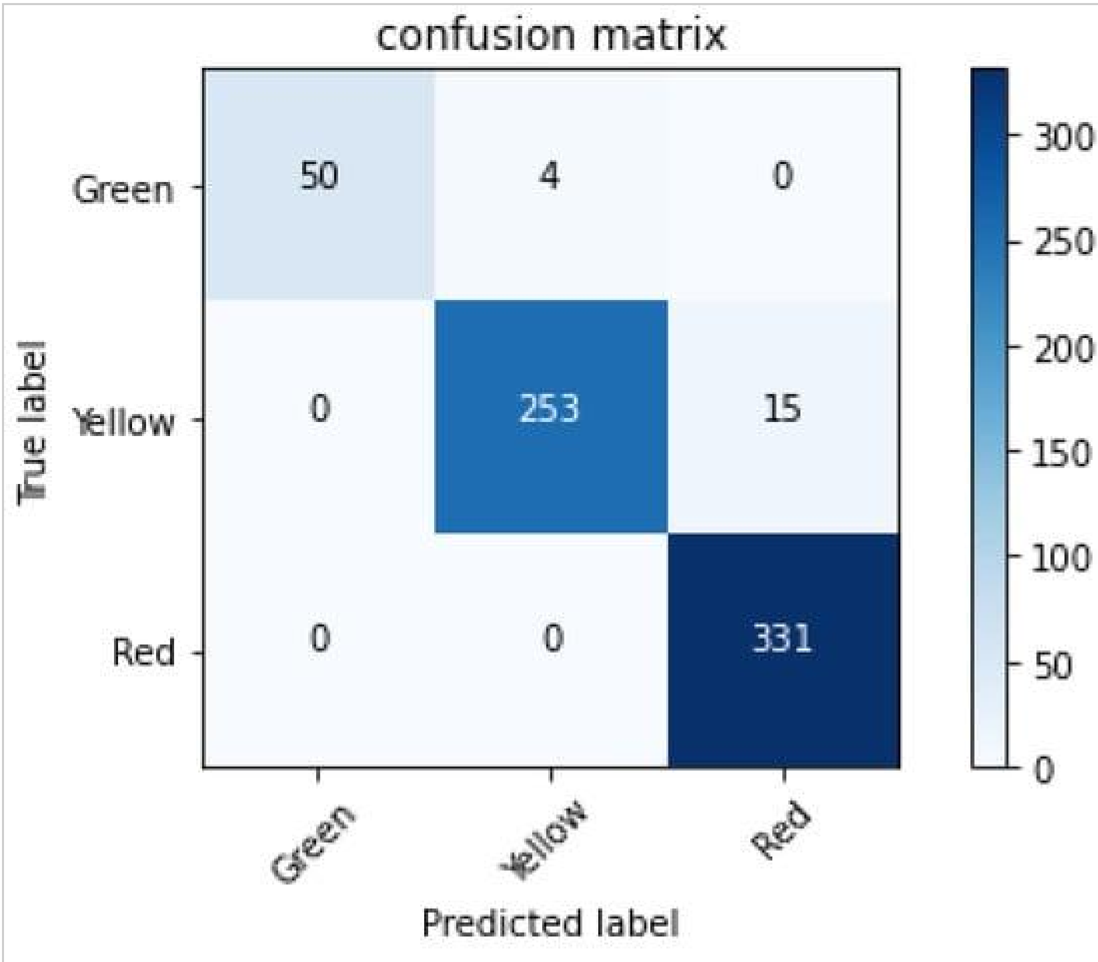



**Figure 11.** ANN confusion matrix on San Lorenzo test case.

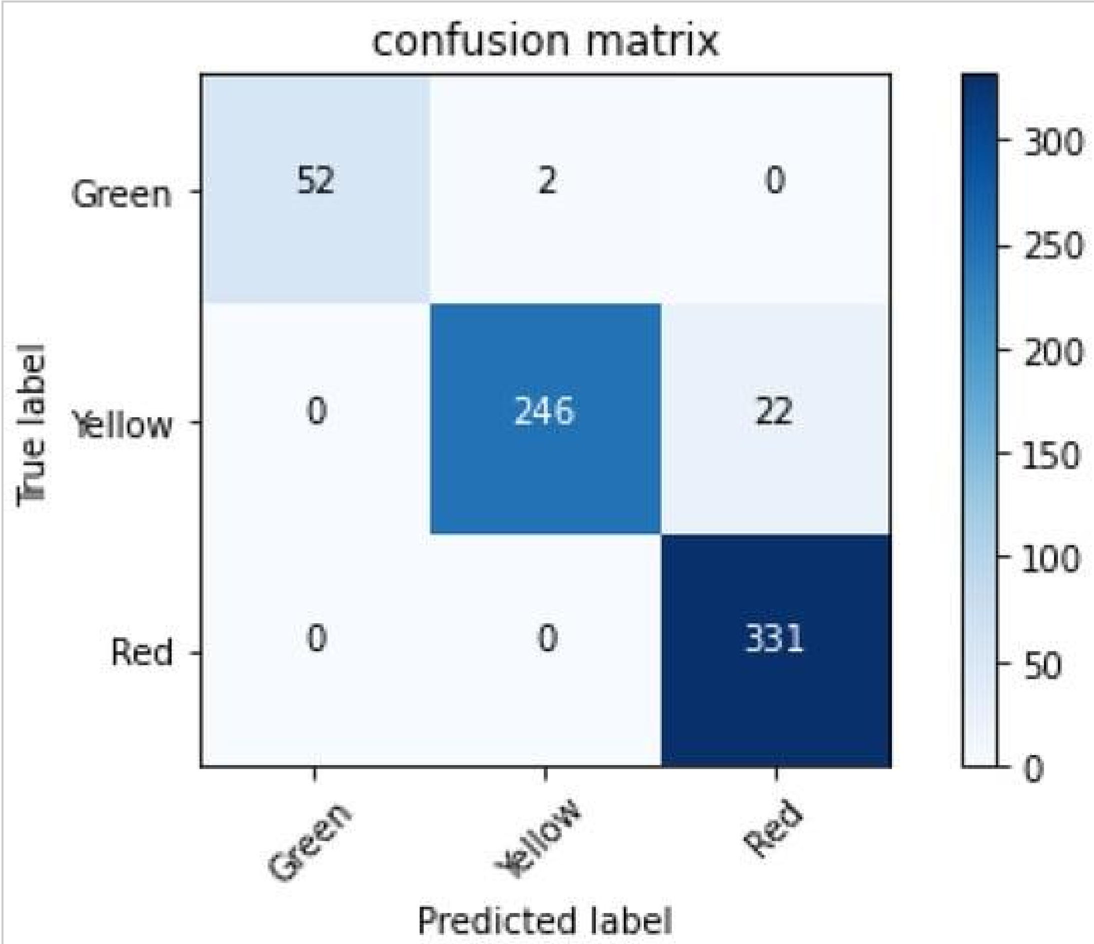



**Figure 12.** KNN confusion matrix on San Lorenzo test case.

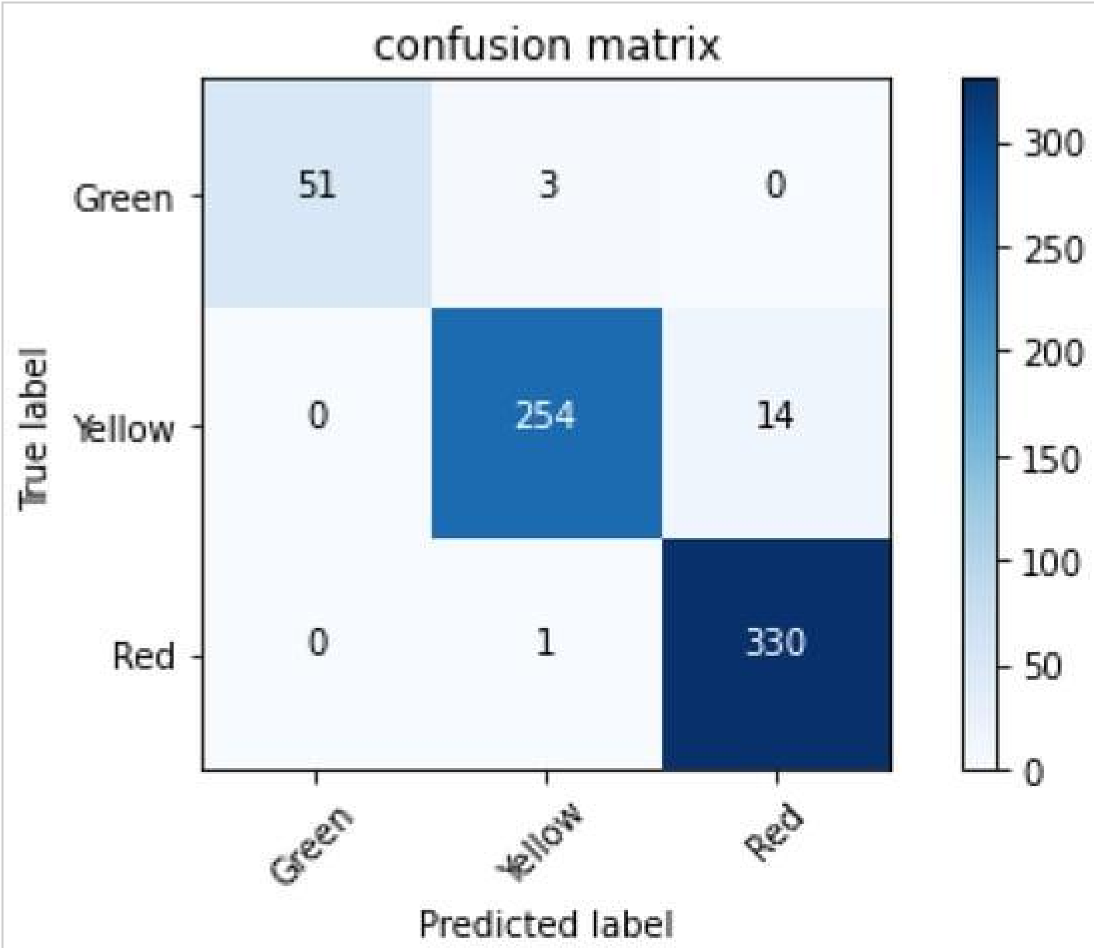



**Figure 13.** Random Forest confusion matrix on San Lorenzo test case.

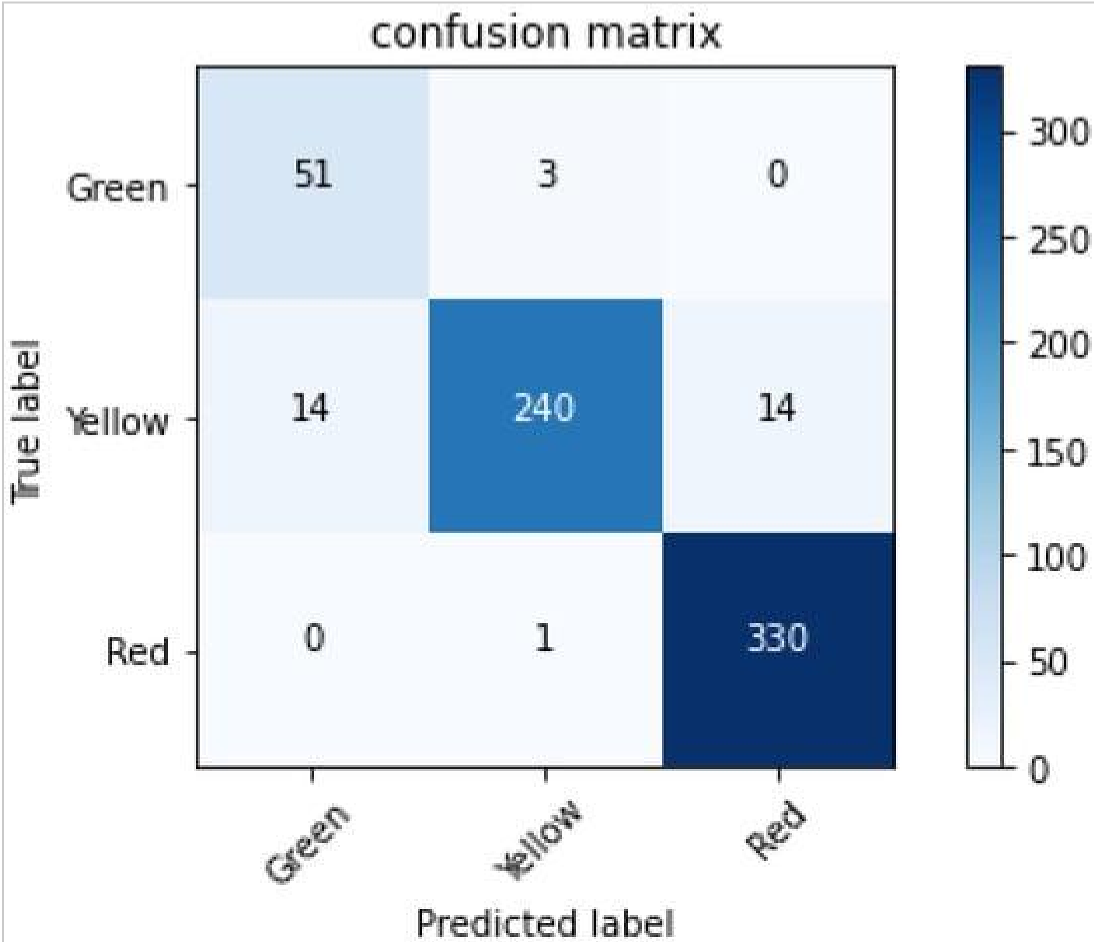



**Figure 14.** Gradient boosting tree confusion matrix on San Lorenzo test case.

As happened during the first training and testing phase, the four considered models obtained very high and very similar results. The metrics adopted to evaluate the models are *accuracy*, *precision*, *recall*, and *f1-score*. The results are shown in **Table 13**, **Table 14**, **Table 15** and **Table 16**.

**Table 13.** ANN metrics on San Lorenzo test case.

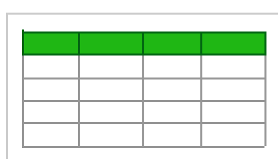


**Table 14.** KNN metrics on San Lorenzo test case.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

**Table 15.** Random forest metrics on San Lorenzo test case.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

**Table 16.** Gradient boosting tree metrics on San Lorenzo test case.

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |

Regarding *accuracy*, the best models are Random Forest and ANN, with 97%, followed by KNN (96%) and GB (95%).

Regarding *precision* in the classification of labels 0, 1 and 2, Random Forest and KNN reach the highest values (1.00, 0.98, 0.96 and 1.00, 0.99 and 0.94, respectively), followed by ANN (0.96, 0.98, 0.96) and GB, with the worst performance in recognizing the 0 label (0.78, 0.98, 0.96).

The *recall* values again see Random Forest and KNN as the best models (all values between 0.92 and 1.00), followed by ANN and GB with a few hundredths of a waste. Finally, even for the *f1-score* values the best models were RF and KNN (0.97–0.98 and 0.95–0.98, respectively), followed by ANN (0.92–0.98) and finally GB (0.86–0.98).

In conclusion, despite the difference between the performances obtained being centesimal, the GB is the worst model in all comparisons. The RF, on the other hand, always stands out as the best result. The ANN, a structurally more sophisticated model, is equal to the RF for accuracy but performs slightly worse according to the other metrics.

Overall, the chosen models performed well in the problem of classifying the level of parking availability of a street segment and confirm the possibility of designing a Parking Availability Classifier that relies on data collectible by smartphone sensors. Compared to other studies that adopted the same evaluation metric (accuracy) but used different sources of data and environments, our models performed better (97% against the 94.37% of [15], the 81% of [16], and the 87.82% of [17]

Conclusion:

The proposed work aimed to correctly classify the parking availability on a given road on a segment level.

First, using the CPS we developed, we generated a large amount of realistic data about cruising, parking, and unparking events in the San Giovanni city area of Rome.

Once sufficient data was generated, some preprocessing steps were applied to build a training dataset for the machine learning models. This phase included applying map-matching techniques to the collected trip data, identifying ten features, and a dataset containing 761 segment samples.

We trained four machine learning models to classify the availability of on-street parking for a given road segment. All models performed well both in the training and testing phases.

To further evaluate the four models, we generated simulated data about 61 segments of a different city area (San Lorenzo, Rome) that differs from San Giovanni as it is a smaller and more traffic-congested neighbor. The results achieved by the models in this test case were very high, reaching 97% accuracy for all the street segments where we collected 30 or more user actions.

As the four models obtained very high results, especially the RF and the ANN, we confirmed that it is possible to design a PAC based on data collectible by smartphones.