Technical report :

In this study, a dataset consisting of KTAS (Korean Triage and Acuity Scales) score data assigned by a nurse and an expert doctor, taking into account the vital signs of patients in emergency departments of hospitals in South Korea, was studied. KTAS is a five-level triage scale developed in 2012 and based on CTAS, modified to apply to the medical situation in South Korea. After the critical first-look assessment, emergency nurses determine the triage level by evaluating the primary considerations.

In the first part of the study, efforts were made to clean the data and make it ready for analysis. Information was obtained about the variables in the data set. The descriptions of the values assigned to the categorical variables were examined. The articles of previous studies on triage in different countries were researched to obtain information about the dataset.

Then, the null values in the data set were determined. The variable that contains too many null values and is difficult to predict based on a different variable has been removed from the dataset to avoid risk. The same method was applied for non-English (Korean) values in the data set and only those values were dropped. For null values in the saturation variable, the most frequently repeating values were determined and assigned to null values.

In the visualization phase of this study, the determinations obtained are as follows:

- The mean age of the participants by gender was 54.89 for female, 53.07 for male.
- Overall, 733 (60%) patients, which was more than half, used a private vehicle.
- The reason for visiting the emergency department was related to non-injurious incidents in 986 (80.8%) cases.
- After treatment in the emergency department, 728 (63.7%) patients were discharged and 361 (2 9.6%) patients were admitted to the wards.
- The comparison of the triage accuracy between the emergency nurses and the experts were:
    - The emergency nurses and experts agreed on a triage score in 1049 (86%) cases.
    - The emergency nurses and experts did not agree on triage scores in 172 (14%) of cases.
    - Among the cases classified as triage, there were 52 (0.04%) cases of over triage (overestimating the severity) and 120 (0.09%) cases of under triage (underestimating the severity).
    - The most frequent reason for mistriage can be the incorrect application of the pain scale score using the KTAS criteria.
    - The second most frequent cause of mistriage can be the misjudgment of physical symptoms associated with the chief complaint.
- There is no significant difference in the number of patient visits per hour for triage accuracy. (mean: 7.49, standard deviation: 3.145)

In the last part of the analysis, the model was developed using the k- means algorithm. K-means value was set to 5 (based on the triage grouping in the data set). The data was splitted two parts as train and test data. Then the model trained and predicted. As a result of this clustering, the predictive power (k-score) of the model was found 48%. We can say that this value is quite low.
We can say that the variables we put into the model for this reason. Variable selection should be made with a specialist about the data and after the expert opinion is received, the process should be done to improve the predictive power of the model, such as hyper parameter tuning.