

IETT and Otobus A.S. are responsible institutions for public bus transportation in Istanbul. People use those buses may encounter some problems about bus, timing, driver, stop etc. So they can report the problems to Istanbul Municipality (IBB) via its “Beyaz Masa” channels. Our data is compiled from this reports from both civilians and personnel of IBB.

Data format that we will work on is Comma Separated Values (CSV) with the columns labelled Durum (result) and Rapor Açıklaması (the text from reporter). There are 64314 records which have 3 different results (Durum) as Menfi, Müspet, Uyarı. The result labels are given to report records as result of the investigation and council decision. Müspet record means no punishment required and there is no guilty, Uyarı record means there is guilty but no punishment given for this time, Menfi means punishment applied to the driver.

In the CSV file we have 48702 Müspet, 14576 Uyarı and 3189 Menfi records. The council want to automate the elimination of Müspet records with a ML algorithm that trained by these records. The auto Müspet labelled records will be filtered before posting to the council.

Since the Rapor Açıklaması came from humans we need to employ some NLP algorithms. When we look at records we see some records have date data and some other not. Some records we saw is about missing bus in a stop, bad behaviour from drivers, smoking driver, full-day working driver, driver passing stop, bus is not convenient for transportation etc.

#	Group Members	Department	Semester	Number
1	İsmail Yavuz	Computer Engineering	8	150115503
2	Arif Taha Ay	Computer Engineering	8	150116079
3	Faruk Furkan Şişman	Computer Engineering	8	150114025
4	Oğuzhan Demirhan	Industrial Engineering	8	150314064
5	Nur Kızıldağ	Industrial Engineering	8	150314056
6	Habib Burak Can Erdönmez	Bio-Engineering	8	150814028

Table 1

#	Feature Data	Description	Type	Target
1	Durum (Result)	Müspet record means no punishment required and there is no guilty, Uyarı record means there is guilty but no punishment given for this time, Menfi means punishment applied to the driver.	Nominal	Yes

Table 2

#	Class Value	# of instances	Corresponding Percentage
1	Müsbet	48702	0.73
2	Menfi	3189	0.048
3	Uyarı	14576	0.22

Table 3