Ezdin Aslancı - 150112022
Kutalmış Coşkun - 150113018

We chose the following feature sets:

1. $\{POS_{t-1}, POS_t\}$
2. $\{POS_{t-2}, POS_{t-1}, POS_t\}$
3. $\{POS_{t-1}, POS_t, POS_{t+1}\}$
4. $\{POS_{t-2}, POS_{t-1}, POS_t, POS_{t+1}\}$
5. $\{POS_{t-2}, POS_{t-1}, POS_t, POS_{t+1}, POS_{t+2}\}$

We tested these features with the following machine learning methods:

- Logistic Regression
- Decision Trees

We conducted experiments with each combination of these two sets. We obtained the following results:

	Precision	Recall	F1 Score
$\{POS_{t-1}, POS_t\}$	0.88	0.89	0.88
$\{POS_{t-2}, POS_{t-1}, POS_t\}$	0.89	0.90	0.89
$\{POS_{t-1}, POS_t, POS_{t+1}\}$	0.91	0.92	0.91
$\{POS_{t-2}, POS_{t-1}, POS_t, POS_{t+1}\}$	0.92	0.92	0.92
$\{POS_{t-2}, POS_{t-1}, POS_t, POS_{t+1}, POS_{t+2}\}$	0.92	0.92	0.92

Table 1: Precision, Recall and F1 Score metrics for each feature set obtained using Logistic Regression.

	Precision	Recall	F1 Score
$\{POS_{t-1}, POS_t\}$	0.89	0.90	0.89
$\{POS_{t-2}, POS_{t-1}, POS_t\}$	0.92	0.92	0.91
$\{POS_{t-1}, POS_t, POS_{t+1}\}$	0.93	0.93	0.93
$\{POS_{t-2}, POS_{t-1}, POS_t, POS_{t+1}\}$	0.96	0.96	0.96
$\{POS_{t-2}, POS_{t-1}, POS_t, POS_{t+1}, POS_{t+2}\}$	0.98	0.98	0.98

Table 2: Precision, Recall and F1 Score metrics for each feature set obtained using Decision Trees.

We also tried using tokens (first column in the data) in feature sets; but with that approach, we think that our model over-fit and we obtained lower scores.

We observed that increasing the number of POS tags in a feature set increases performance.

Although 0.98 F1 Score looked suspicious to us, using Decision Trees with different feature sets and using the same feature set with a different machine learning algorithm resulted worse so we got sure that we did not cheat.

Here are graphs showing performance of each combination.

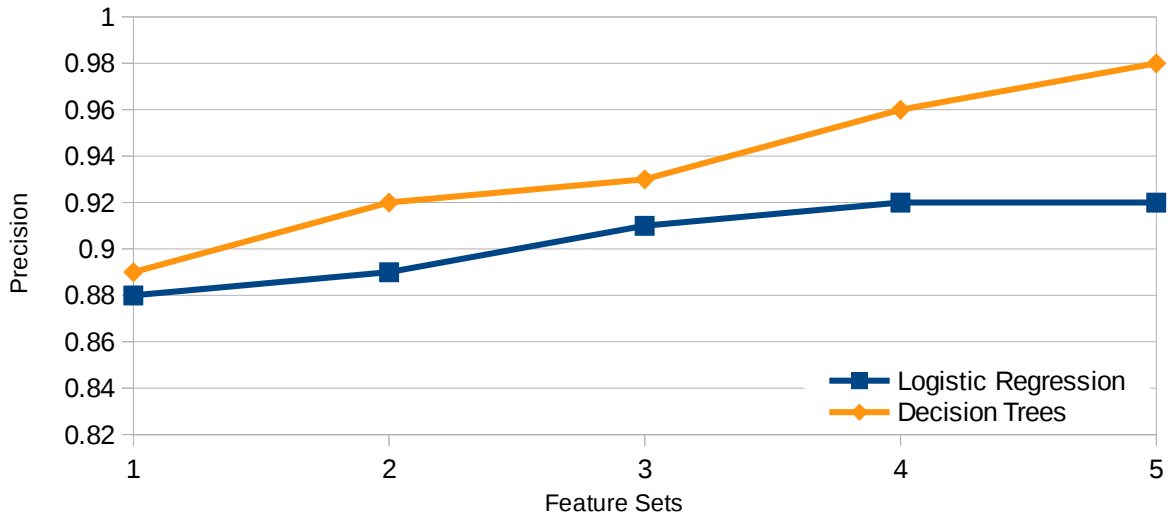

Figure 1: Precision values of feature sets.

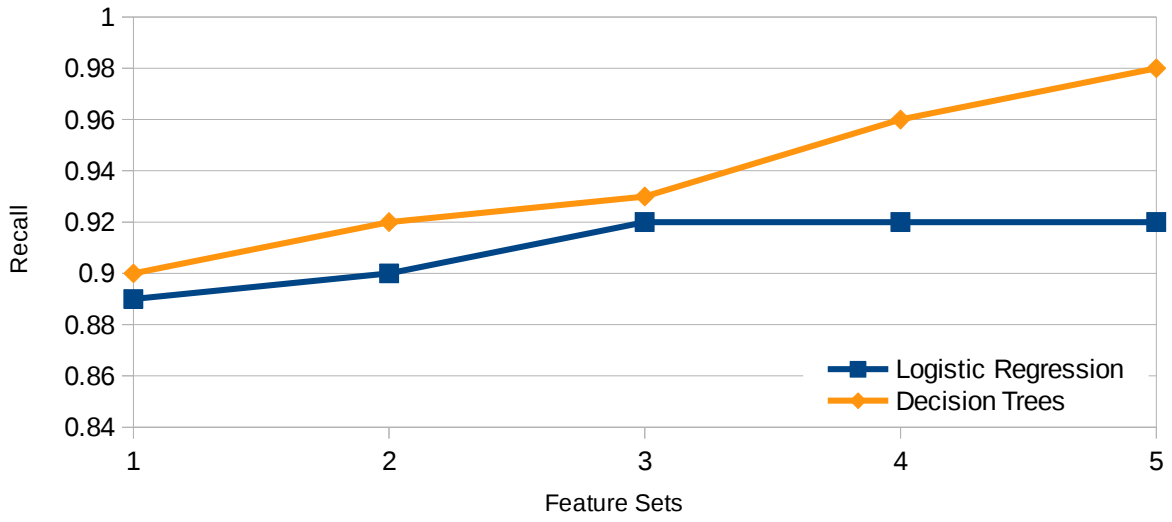

Figure 2: Recall values of feature sets.

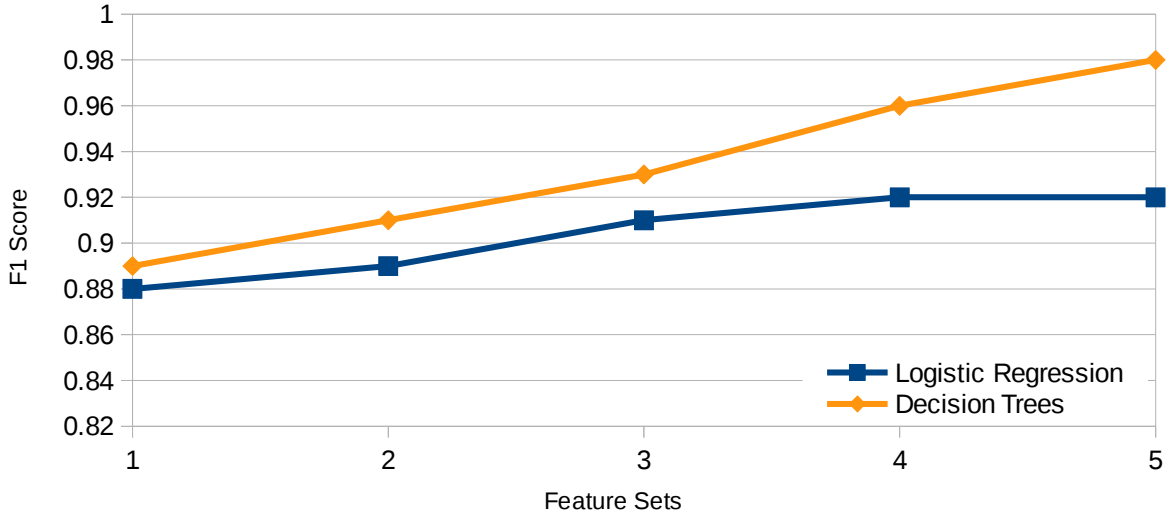

Figure 3: F1 Scores of feature sets.