

PREDICTION MODELS BASED ON MAX-STEMS (or harnessing imbalanced data)

Episode Four: Advanced Examinations

Ahmet Furkan EMREHAN

(matahmet@gmail.com)

EMREHAN

11/17/2021

1



PREDICTION MODELS BASED ON MAX-STEMS

- ▶ Episode One: One-Word Based
- ▶ Episode Two: A Combinatorial Approach
- ▶ Episode Three: Effect of Hyperparameters
- ▶ Episode Four: Advanced Examinations

EMREHAN

11/17/2021

2



INTRODUCTION

- In this chapter, I examine some cases:
- A) How do Accuracy Rates of Categories change, if the sample is changed ?
- B) Let's compare all methods for each category
- C) Can a Trigonometric Approach contribute to hyperparametric model ?

EMREHAN

11/17/2021

3



A) How do Accuracy Rates of Categories change, if the sample is changed ?

(Based on Predict Model 1 (On Word))

Accuracy Rates of Categories					
Sample	DÜNYA	SPOR	SANAT	Teknoloji	
1	0.791	0.167	0.029	0.013	
2	0.798	0.166	0.025	0.012	
3	0.783	0.181	0.023	0.013	
4	0.794	0.169	0.024	0.013	
5	0.792	0.172	0.024	0.012	
6	0.793	0.175	0.024	0.009	
7	0.797	0.172	0.022	0.010	
8	0.809	0.159	0.022	0.010	
9	0.793	0.175	0.024	0.009	
10	0.797	0.172	0.022	0.010	

Result: Stable against samples

Note: Naturally Samples intersect wit each other at %80 level.

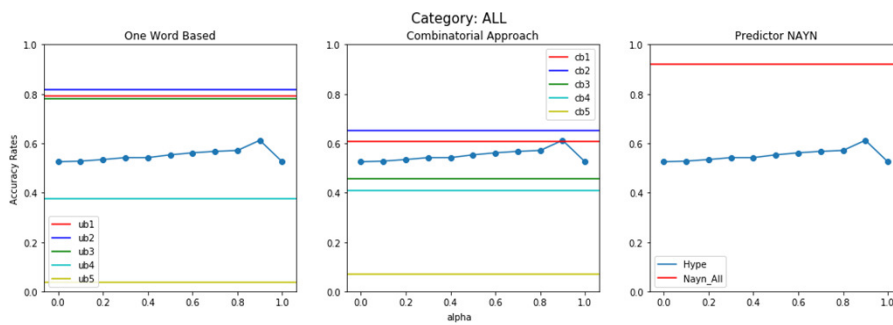
EMREHAN

11/17/2021

4



B) Let's compare all methods for each category



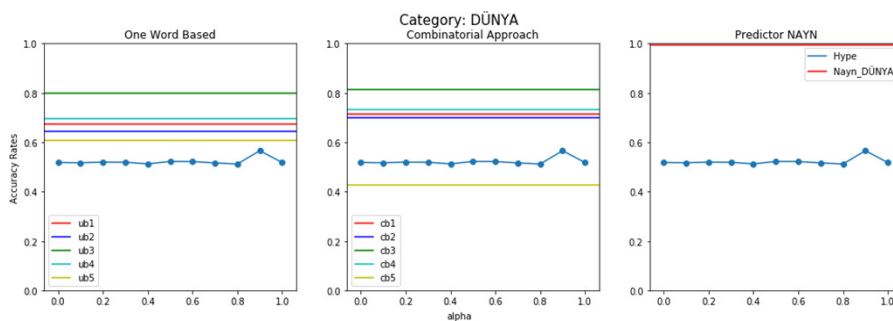
EMREHAN

11/17/2021

5



B) Let's compare all methods for each category



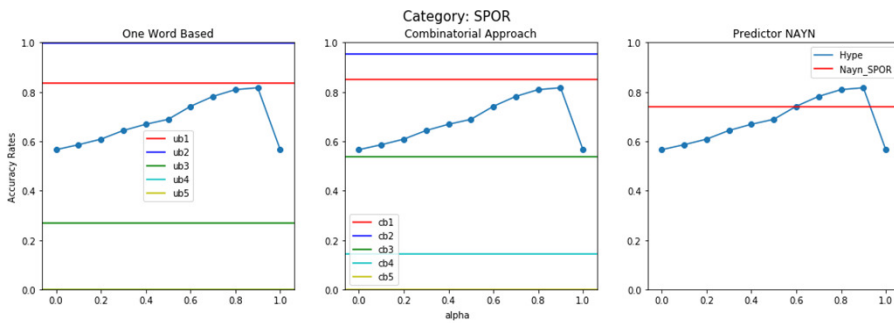
EMREHAN

11/17/2021

6



B) Let's compare all methods for each category



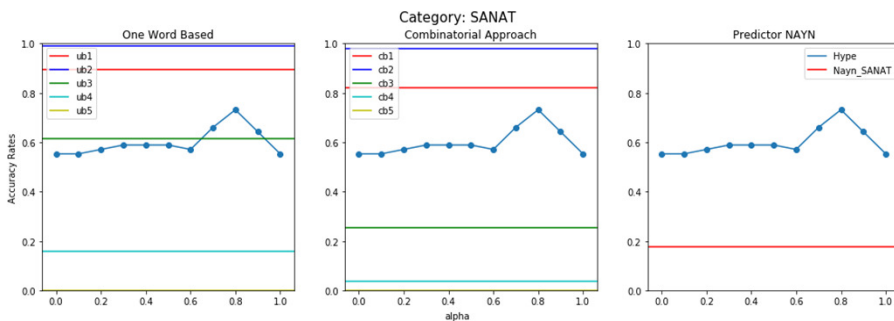
EMREHAN

11/17/2021

7



B) Let's compare all methods for each category



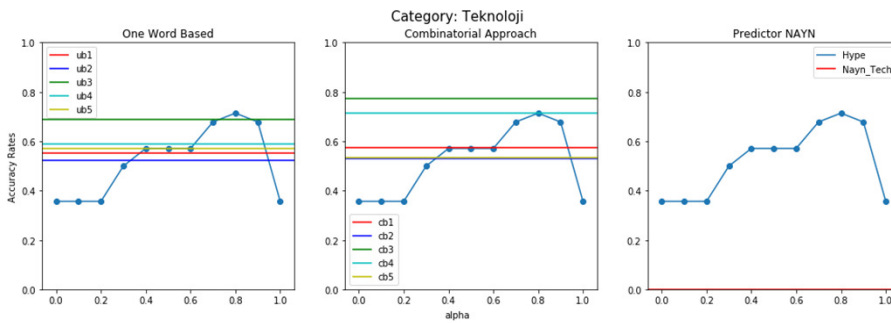
EMREHAN

11/17/2021

8



B) Let's compare all methods for each category



EMREHAN

11/17/2021

9



B) Let's compare all methods for each category

- Conclusion:
- Whereas NAYN predictor is successful at predicting documents labelled with «DÜNYA», other prediction models (one stems and combinatorial approach) is more successful at predicting documents labelled with other categories.

EMREHAN

11/17/2021

10



C) Can a Trigonometric Approach contribute to hyperparametric model ?

- In this section we use trigonometric approach to hyperparameters such that $\alpha = \sin(\theta)$ and $\beta = \cos(\theta)$ where $\theta \in \left[0, \frac{\pi}{2}\right]$

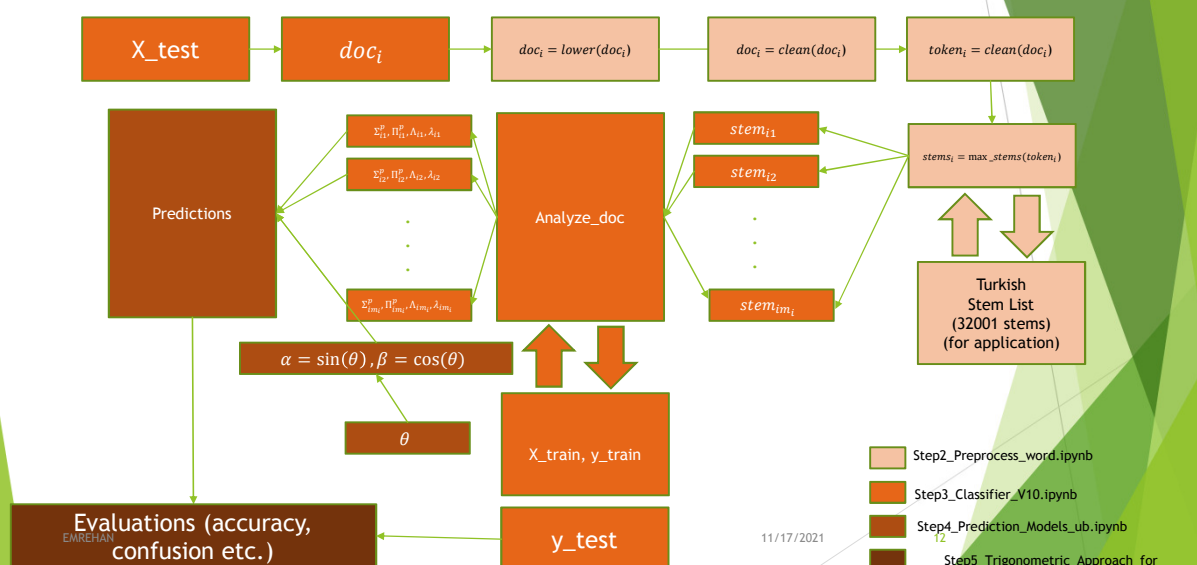
EMREHAN

11/17/2021

11



General Scheme for Application of Trigonometric Approach to Prediction Models with hyperparameter

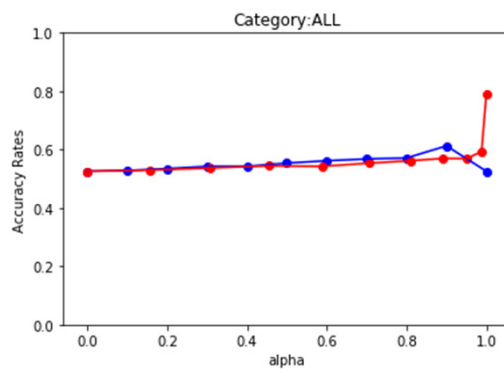


11/17/2021



C) Can a Trigonometric Approach contribute to hyperparametric model ?

► Results



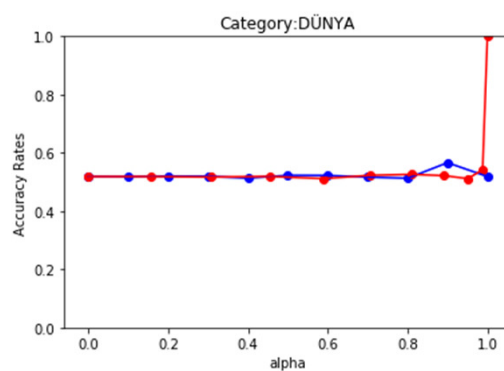
EMREHAN

11/17/2021

13



C) Can a Trigonometric Approach contribute to hyperparametric model ?



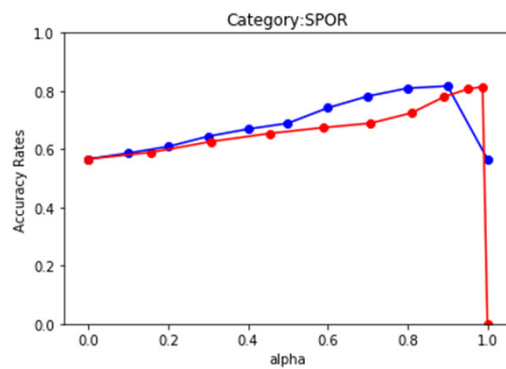
EMREHAN

11/17/2021

14



C) Can a Trigonometric Approach contribute to hyperparametric model ?



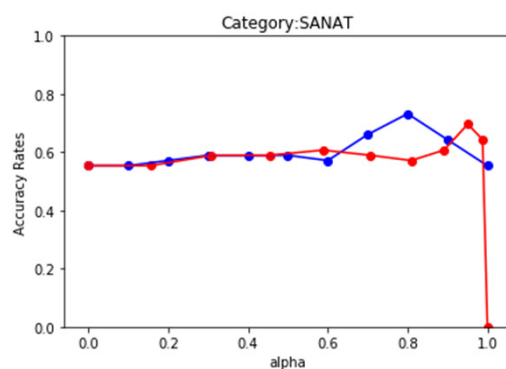
EMREHAN

11/17/2021

15



C) Can a Trigonometric Approach contribute to hyperparametric model ?



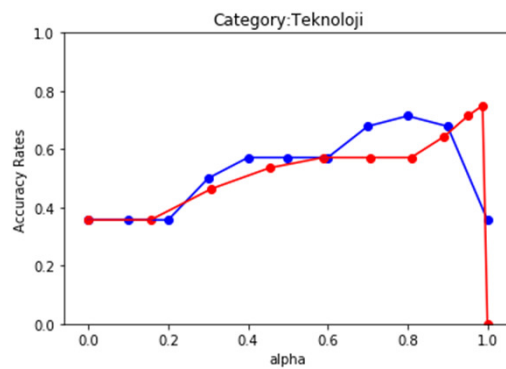
EMREHAN

11/17/2021

16



C) Can a Trigonometric Approach contribute to hyperparametric model ?



EMREHAN

11/17/2021

17



C) Can a Trigonometric Approach contribute to hyperparametric model ?

- Conclusion:
- 1) When $\theta = \frac{\pi}{2} \rightarrow \alpha = 0, \beta = 1$, model produces only «DÜNYA» label, then specificity fails. In other saying, predictions of other categories in test set become «DÜNYA»
- 2) Generally, performance of model based trigonometric functions falls behind that of model based uniform distribution in almost all categories. Only at prediction of documents labelled «Teknoloji», performance of trigonometric approach is a notch better than uniform approach.

EMREHAN

11/17/2021

18



