

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

Episode Three: Effect of Hyperparameters

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

- The model in this section is with hyperparameter.  $\alpha$  and  $\beta$  are the hyperparameters controlling length of max-stem and frequencies of max-stem in main categories respectively.

EMREHAN

11/17/2021

3

## COMPONENTS OF MODELS

- $p$ : index of categories (or labels)
- $Label^p$ : category with  $p$  index
- $n$ : counts of categories (or labels)
- $doc_i$ : document, in test set, with index  $i$  as a sentences or just a headline
- $stem_{ij}$ : stem with index  $j$  of  $doc_i$   
(stem can be chosen as max – stem mentioned previous slides.)
- $m_i$ : counts of  $stem_{ij}$
- $\Sigma^p$ : counts of documents labelled with category with index  $p$  in train set
- $\Sigma_{ij}^p$ : counts of documents, which include  $stem_{ij}$ , labelled with category with index  $p$  in train set

EMREHAN

11/17/2021

4

## COMPONENTS OF MODELS

- ▶  $\Lambda_{ij} := \text{Label}^q$  where  $q = \arg \max_p \Sigma_{ij}^p$
- ▶  $\Lambda_i^p$ : counts of  $\Lambda_{ij}$  which equals to  $\text{Label}^p$
- ▶  $\lambda_{ij}$ : length of  $\text{stem}_{ij}$
- ▶  $\rho_i^p := \frac{\sum_{j=1}^{m_i} \Sigma_{ij}^p}{\Sigma^p} *$
- ▶ \* in case that  $\Sigma^p = 0, \rho_i^p := 0$
- ▶  $\Pi_{ij}^p := \frac{\Sigma_{ij}^p}{\sum_{q=1}^n \Sigma_{ij}^q}$   
(it can be considered as probability of  $\text{stem}_{ij}$  labelled with category with  $p$  index)

EMREHAN

11/17/2021

5



## COMPONENTS OF MODELS

- ▶  $\overline{\Pi}_i^p := \text{average}_{j*} (\Pi_{ij*}^p)$  such that all "j\*"s meet the condition  $\Pi_{ij*}^p > 0$   
\* in case that  $\Sigma_{ij}^p = 0$  for all  $p = 1, 2, \dots, n, \overline{\Pi}_i^p = 0$
- ▶  $\widehat{\Pi}_i^p := \max_j (\Pi_{ij}^p)$

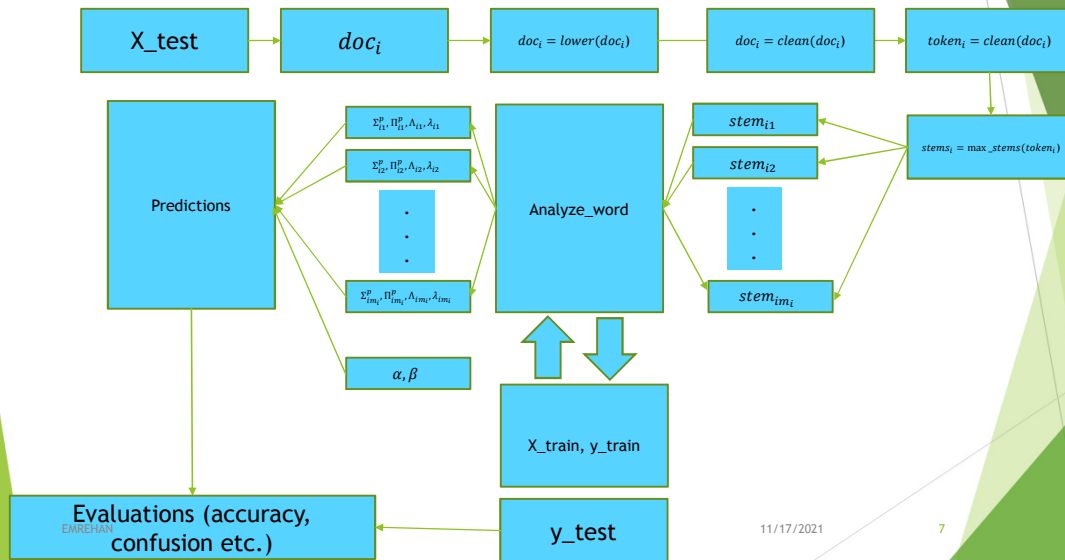
EMREHAN

11/17/2021

6



## General Scheme for Prediction Models with hyperparameters



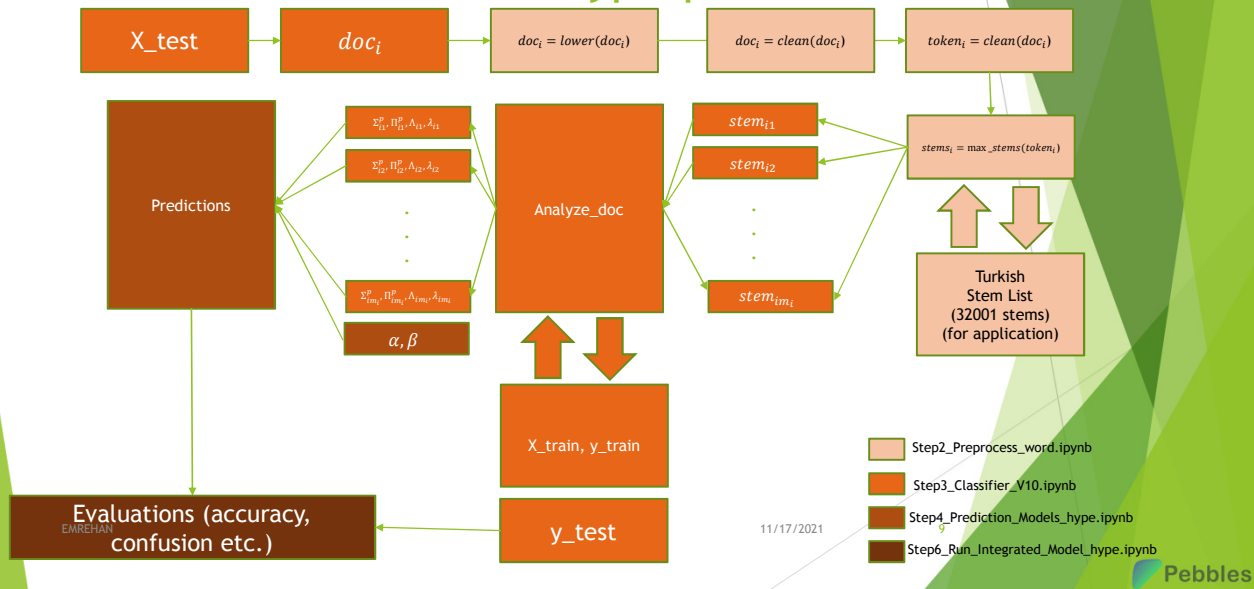
11/17/2021

7

## Model with hyperparameter

- $predict_{type}(doc_i, \alpha, \beta) = Label^q$  where  $q = \arg \max_p (\max_j [(\lambda_{ij})^\alpha * (\pi_{ij}^p)^\beta])$

## General Scheme for Application of Prediction Models with hyperparameter

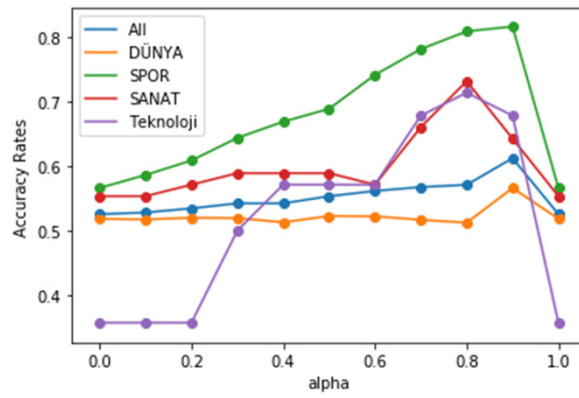


## Results

- Accuracy Rates for All Categories and sub-categories over  $\alpha$  and  $\beta$  (or  $1 - \alpha$ ) in  $[0,1] \times [0,1]$

alpha	beta	All	DÜNYA	SPOR	SANAT	Teknoloji
0	1	0.5256	0.5185	0.5664	0.5536	0.3571
0.1	0.9	0.5282	0.5174	0.5865	0.5536	0.3571
0.2	0.8	0.5346	0.5201	0.609	0.5714	0.3571
0.3	0.7	0.5424	0.5195	0.6441	0.5893	0.5
0.4	0.6	0.5424	0.513	0.6692	0.5893	0.5714
0.5	0.5	0.5535	0.5228	0.6892	0.5893	0.5714
0.6	0.4	0.5617	0.5223	0.7419	0.5714	0.5714
0.7	0.3	0.5677	0.5168	0.782	0.6607	0.6786
0.8	0.2	0.5712	0.5125	0.8095	0.7321	0.7143
0.9	0.1	0.6125	0.5662	0.817	0.6429	0.6786
1	0	0.5256	0.5185	0.5664	0.5536	0.3571

## Results



EMREHAN

11/17/2021

11