Towards Suicide Prevention:

Early detection of depression on social media

Víctor Leiva & Ana Freire

upf. Universitat Pompeu Fabra Barcelona

Unitat de Coordinació Acadèmica d'Enginyeries i Tecnologies de la Informació i les Comunicacions

EXCELENCIA MARÍA DE MAEZTU

ana.freire@upf.edu

@ana__freire

Motivation

800,000
people die due to suicide every year

40
seconds a new death by suicide

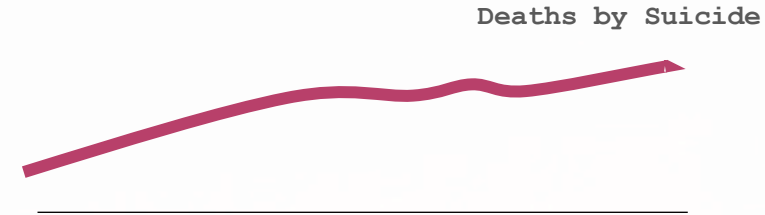
2nd
leading cause of death among 15–29-year-olds

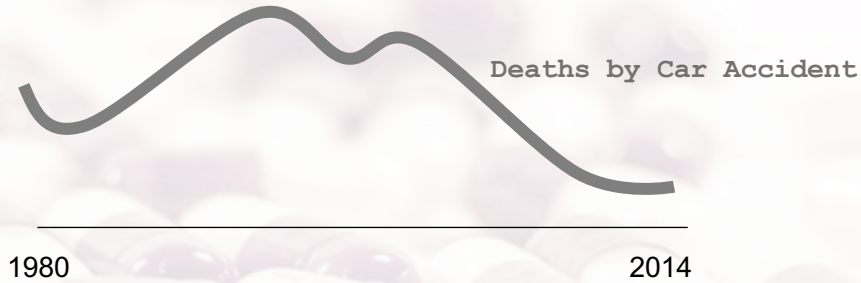
78%
global suicides occur in low- and middle-income countries

World Health Organization

Motivation


Deaths by Suicide
3910
Spain, 2014
Deaths by Car accident
1873


Deaths by Suicide
1980
2014


Deaths by Car Accident
1980
2014

Motivation

90% of suicides can be attributed to mental illnesses

Depression is the leading one

Background

Social media and depression

Depressive Moods of Users Portrayed in Twitter (Park,2012)

- People that suffers from mental disease tend to post more about it.

A content analysis of depression-related Tweets (Cavazos-Rehg, 2016)

- Tweet content could help to identify depression

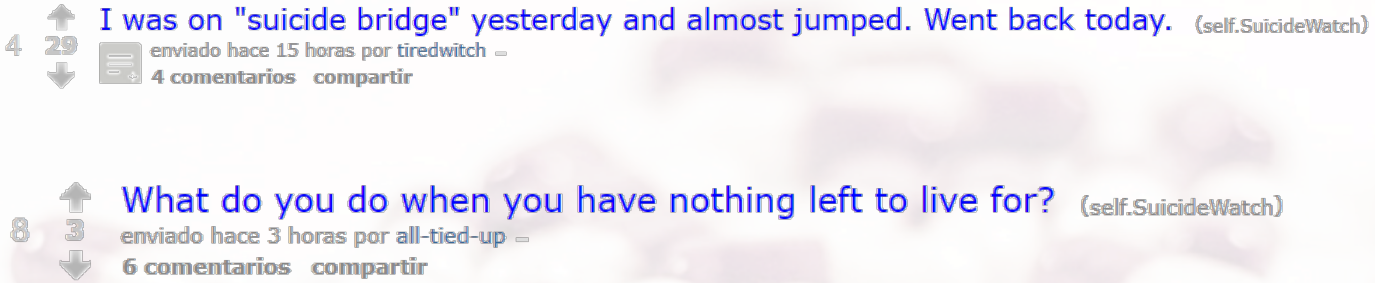
Social media as a measurement tool of depression in populations (De Choudhury, 2013)

…

Background

Early Risk Detection

A Test Collection for Research on Depression and Language Use (Losada, 2016)



Background

Early Risk Detection

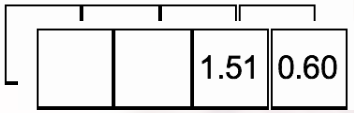
A Test Collection for Research on Depression and Language Use (Losada, 2016)

Messages from depressive users

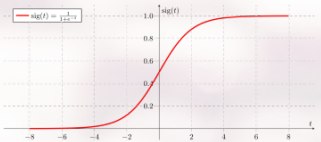
reddit

Messages from NON-depressive users

TF*IDF vectorization



Logistic Regression



Evaluation

- Precision
- Recall
- F-measure
- ERDE

Background

Early Risk Detection

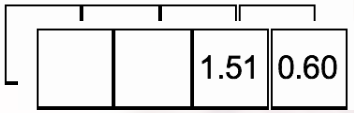
A Test Collection for Research on Depression and Language Use (Losada, 2016)

reddit

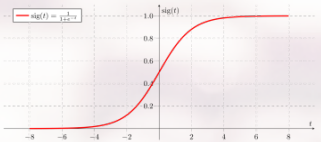
Messages from depressive users

Messages from NON-depressive users

TF*IDF vectorization



Logistic Regression



Evaluation

- Precision
- Recall
- F-measure
- ERDE

Background

Early Risk Detection

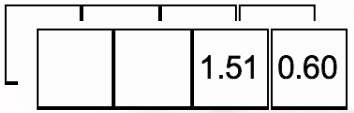
A Test Collection for Research on Depression and Language Use (Losada, 2016)

Messages from depressive users

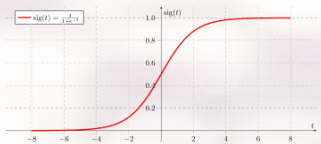
reddit

Messages from NON-depressive users

TF*IDF vectorization



Logistic Regression



Evaluation

- Precision
- Recall
- F-measure
- ERDE

Proposal

Propose a *new model* for early-risk detection of depression by:

- **Extracting** text **features** in a more accurate way.

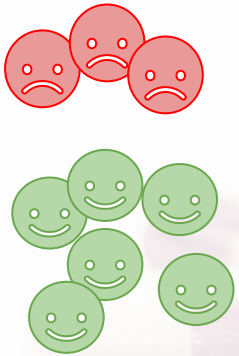
Proposal

Propose a *new model* for early-risk detection of depression by:

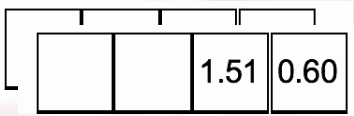
- **Extracting** text **features** in a more accurate way.

+ Dimensionality Reduction
+ Sentiment Analysis

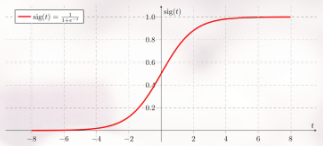
reddit



TF*IDF vectorization



Logistic Regression



Evaluation

- Precision
- Recall
- F-measure
- ERDE

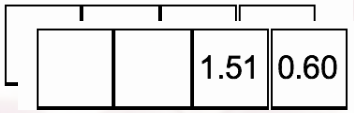
Proposal

Propose a *new model* for early-risk detection of depression by:

- **Exploring** the behaviour of **classic** and **modern** machine learning techniques
- Studying if **genetic algorithms** can help in detecting depression in social media

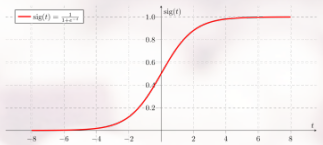
reddit

TF*IDF vectorization



+ KNN, SVM, RF, Soft-Voting
+ Parameter optimization (genetic alg.)

Logistic Regression



Evaluation

- Precision
- Recall
- F-measure
- ERDE

Dataset

A Test Collection for Research on Depression and Language Use (Losada, 2016)

125 depressed users

752 non-depressed users

messages/user: 10 - 2000 (average = 607)

Dataset

- **Training set:**

83 users

403 users

messages/user: 10 - 2000 (average = 607)

- **Evaluation set:**

52 users

349 users

Data pre-processing (1/3)

TF*IDF vectorization

$$w_{i,j} = \mathrm{tf}_{i,j} \cdot \log \frac{N}{n_i}$$

$w_{i,j}$ weight assigned to term *i* in document *j*

$\mathrm{tf}_{i,j}$ number of occurrence of term *i* in document *j*

N number of documents in entire collection

n_i number of documents with term *i*

$W_{i,j}$

	Message 1	2	3	4
complicated			1.51	0.60
sadness	0.50	0.13	0.38	
me	0.63		0.50	0.38
great				
interesting		0.60		
depression	0.90		2.11	
boyfriend		0.75	0.13	0.50
…	1.20			

Data pre-processing (1/3)

TF*IDF vectorization

$$w_{i,j} = \mathrm{tf}_{i,j} \cdot \log \frac{N}{n_i}$$

$w_{i,j}$ weight assigned to term *i* in document *j*

$\mathrm{tf}_{i,j}$ number of occurrence of term *i* in document *j*

N number of documents in entire collection

n_i number of documents with term *i*

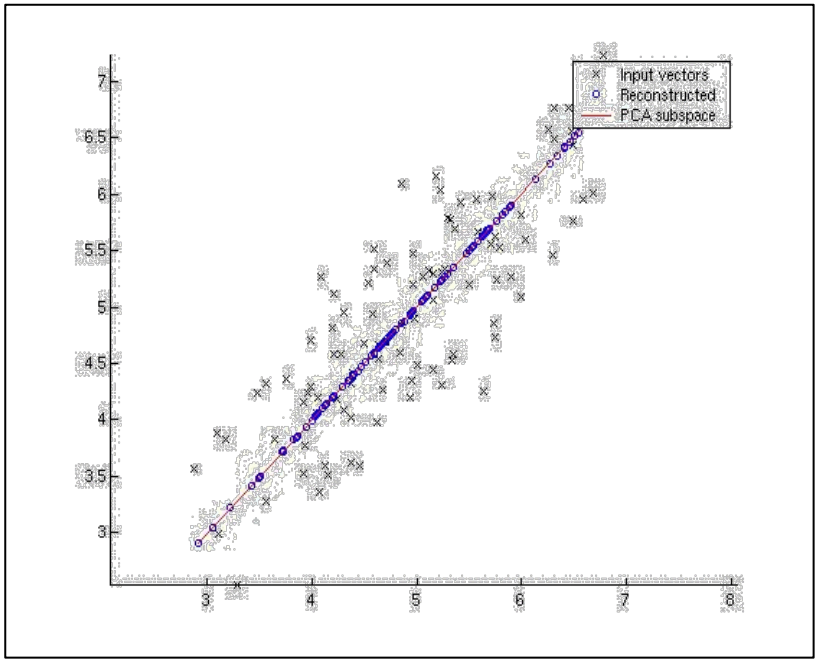
$W_{i,j}$

	Message 1	2	3	4
complicated			1.51	0.60
sadness	0.50	0.13	0.38	
me	0.63		0.50	0.38
great				
interesting		0.60		
depression	0.90		2.11	
boyfriend		0.75	0.13	0.50
…	1.20			

Almost 13k features!!

Data pre-processing (2/3)

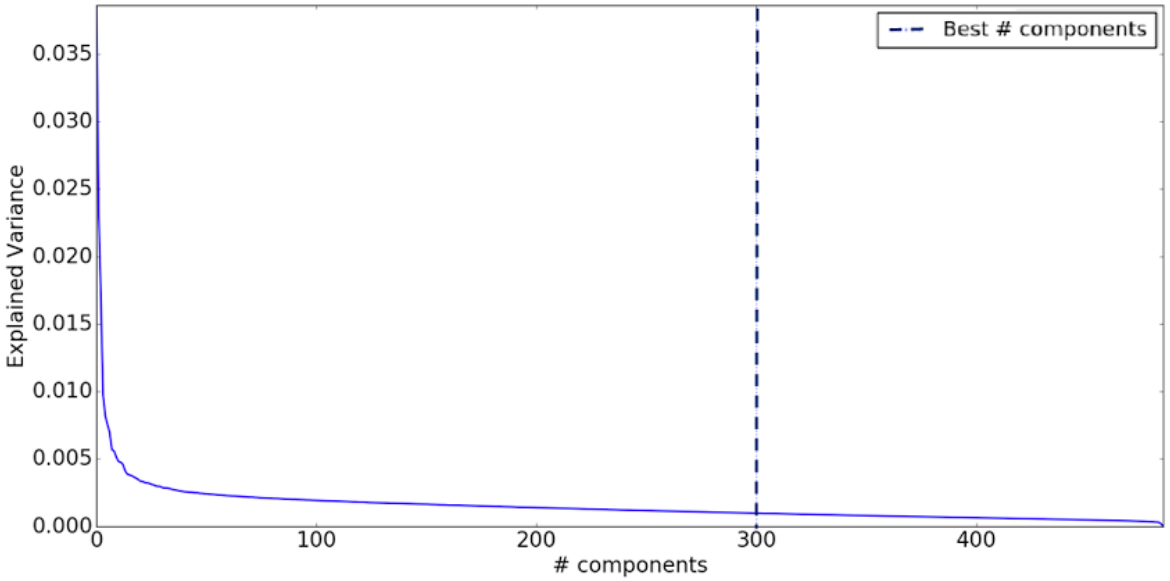
Principal Component Analysis (PCA)



Principal Component Analysis used to find the 1D representation of the input 2D data with the minimal reconstruction error.

Data pre-processing (2/3)

Principal Component Analysis (PCA)



Almost 13k features!!

300

Data pre-processing (3/3)

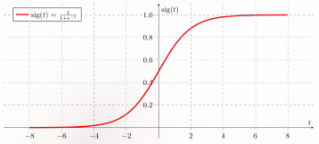
Adding extra features

- Valence Aware Dictionary and sEntiment Reasoner (VADER)

```
Make sure you :) or :D today!---------------------------------- {'neg': 0.0, 'neu': 0.294, 'pos': 0.706}
Today SUX!----------------------------------------------------- {'neg': 0.779, 'neu': 0.221, 'pos': 0.0}
Today only kinda sux! But I'll get by, lol----------------------- {'neg': 0.179, 'neu': 0.569, 'pos': 0.251}
```

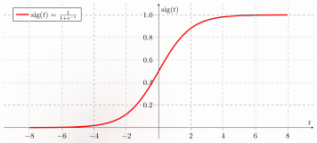
Machine learning algorithms

Logistic Regression:

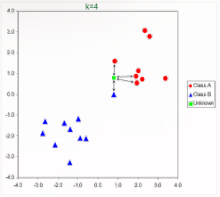


Machine learning algorithms

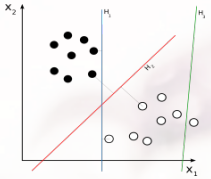
Logistic Regression:



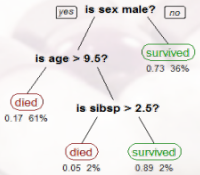
K - Nearest Neighbours:



Support Vector Machine:

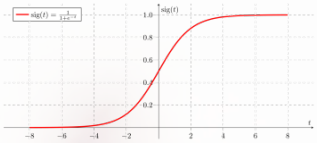


Random Forests:

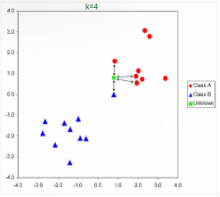


Machine learning algorithms

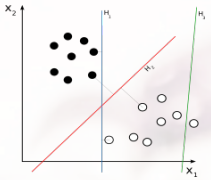
Logistic Regression:



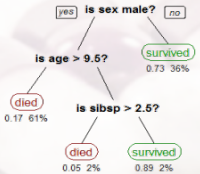
K - Nearest Neighbours:



Support Vector Machine:



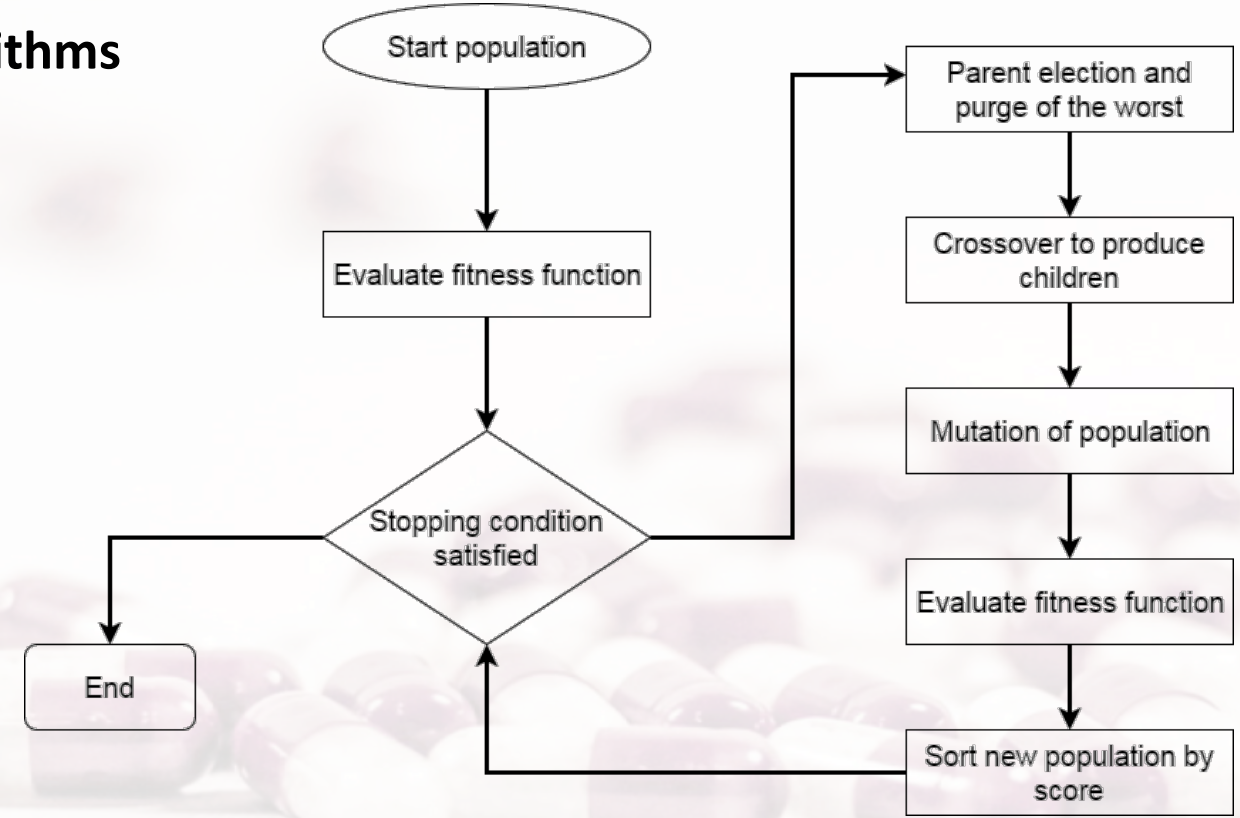
Random Forests:



Voting algorithm

Parameter optimization (for the voting algorithm)

Genetic algorithms

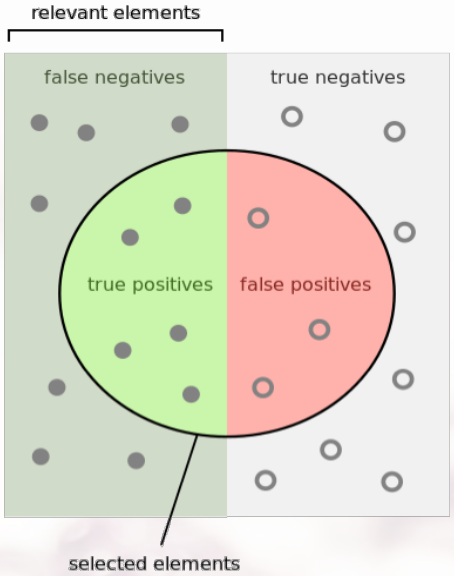


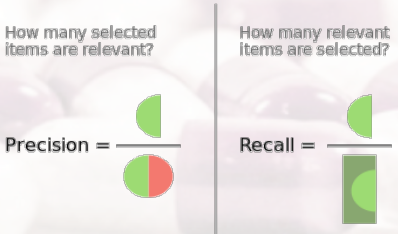
Evaluation

Precision

Recall

F-measure



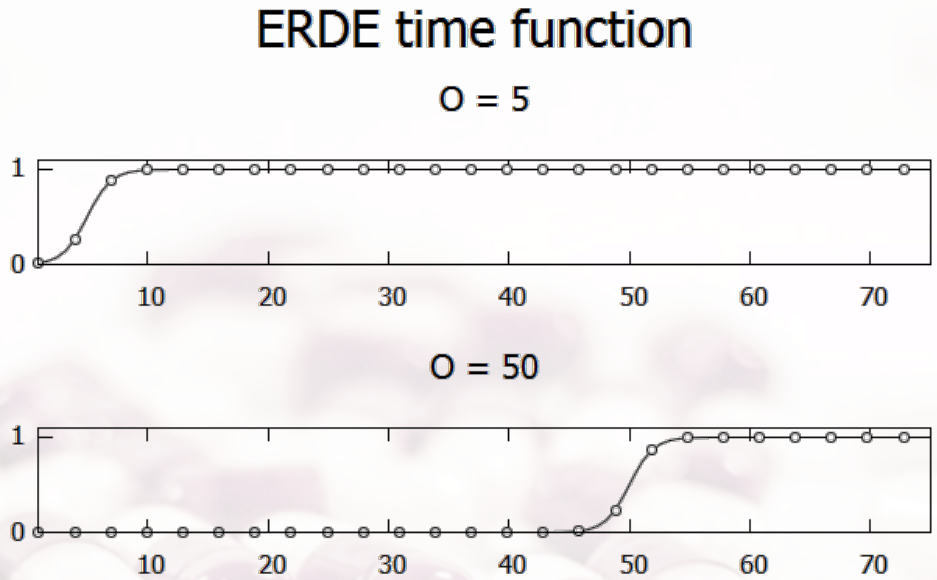


Evaluation

Early Risk Detection Error

$$ERDE_O(k) = \begin{cases} c_{fp} & if\ FP \\ c_{fn} & if\ FN \\ f_O(k)\ c_{tp} & if\ TP \\ 0 & if\ TN \end{cases}$$

$$f_O(k) = 1 - \frac{1}{1 + e^{k-O}}$$



Experimental pipeline

1. Determine **baselines**
2. Use different **ML methods** and build a Soft-Voting Algorithm
3. Refine **features**:
    a. dimensionality reduction (**PCA**)
    b. adding new features (**VADER**)
4. Parameter optimization through **Genetic Algorithms**

Baselines

		ERDE5	ERDE50	P	R	F1
Minority		0.113	0.113	0.13	**1.00**	0.30
Random		0.130	0.130	0.11	0.46	0.18
First	10	0.100	0.098	**0.63**	0.29	0.39
	50	0.083	0.081	0.57	0.46	0.51
	100	0.074	0.072	0.55	0.56	0.55
Logistic Regression	0.5	**0.058**	**0.052**	0.41	0.77	0.54
	0.75	0.072	0.069	0.61	0.54	**0.57**

Choosing the best baseline from the ones proposed in:

A Test Collection for Research on Depression and Language Use (Losada, 2016)

B

Learning Algorithm

		ERDE5	ERDE50	P	R	F1
Logistic Regression	0.5	**0.058**	**0.052**	0.41	0.77	0.54
Support Vector Machines	0.5	0.070	0.065	0.58	0.58	0.58
	0.75	0.091	0.089	0.78	0.35	0.48
Random Forest	0.5	0.083	0.075	0.22	0.83	0.35
	0.75	0.111	0.110	0.82	0.17	0.29
K - Nearest Neighbours	0.5	0.071	0.066	0.23	**0.88**	0.37
	0.75	0.077	0.072	0.59	0.50	0.54
Voting Algorithm	0.5	**0.058**	0.053	0.53	0.69	**0.60**
	0.75	0.091	0.088	**0.86**	0.35	0.49

TF-IDF features (almost 13k)

		ERDE5	ERDE50	P	R	F1
B	Logistic Regression	0.058	**0.052**	0.41	0.77	0.54
LA	Voting Algorithm	0.058	0.053	0.53	0.69	0.60
LA + PCA	Logistic Regression	0.065	0.060	0.49	0.65	0.56
	K - Nearest Neighbours	**0.057**	**0.052**	0.37	0.79	0.51
	Support Vector Machine	0.069	0.064	**0.59**	0.58	0.58
	Random Forest	0.106	0.099	0.15	**0.96**	0.26
	Voting Algorithm	0.062	0.056	0.58	0.65	**0.61**

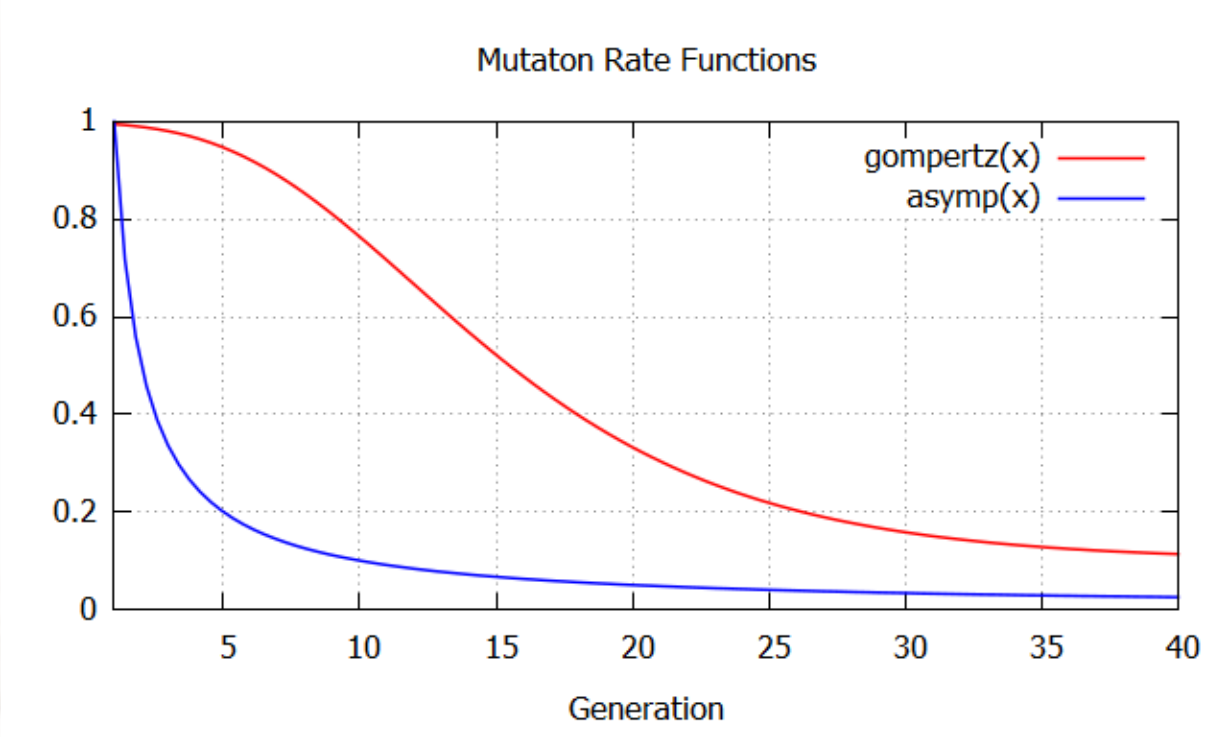
Reduced features (300)

		ERDE5	ERDE50	P	R	F1
B	Logistic Regression	0.058	**0.052**	0.41	0.77	0.54
LA	Voting Algorithm	0.058	0.053	0.50	0.69	0.60
LA + PCA	K - Nearest Neighbours	**0.057**	**0.052**	0.37	0.79	0.51
LA + PCA	Voting Algorithm	0.062	0.056	**0.58**	0.65	**0.61**
LA + PCA + VADER	Logistic Regression	0.063	0.058	0.49	0.67	0.57
LA + PCA + VADER	K - Nearest Neighbours	0.058	0.053	0.38	0.77	0.51
LA + PCA + VADER	Support Vector Machine	0.072	0.066	**0.58**	0.56	0.57
LA + PCA + VADER	Random Forest	0.085	0.078	0.19	**0.96**	0.31
LA + PCA + VADER	Voting Algorithm	0.061	0.056	0.53	0.67	0.59

Reduced features (300) + sentiment features (VADER)

Genetic Algorithms

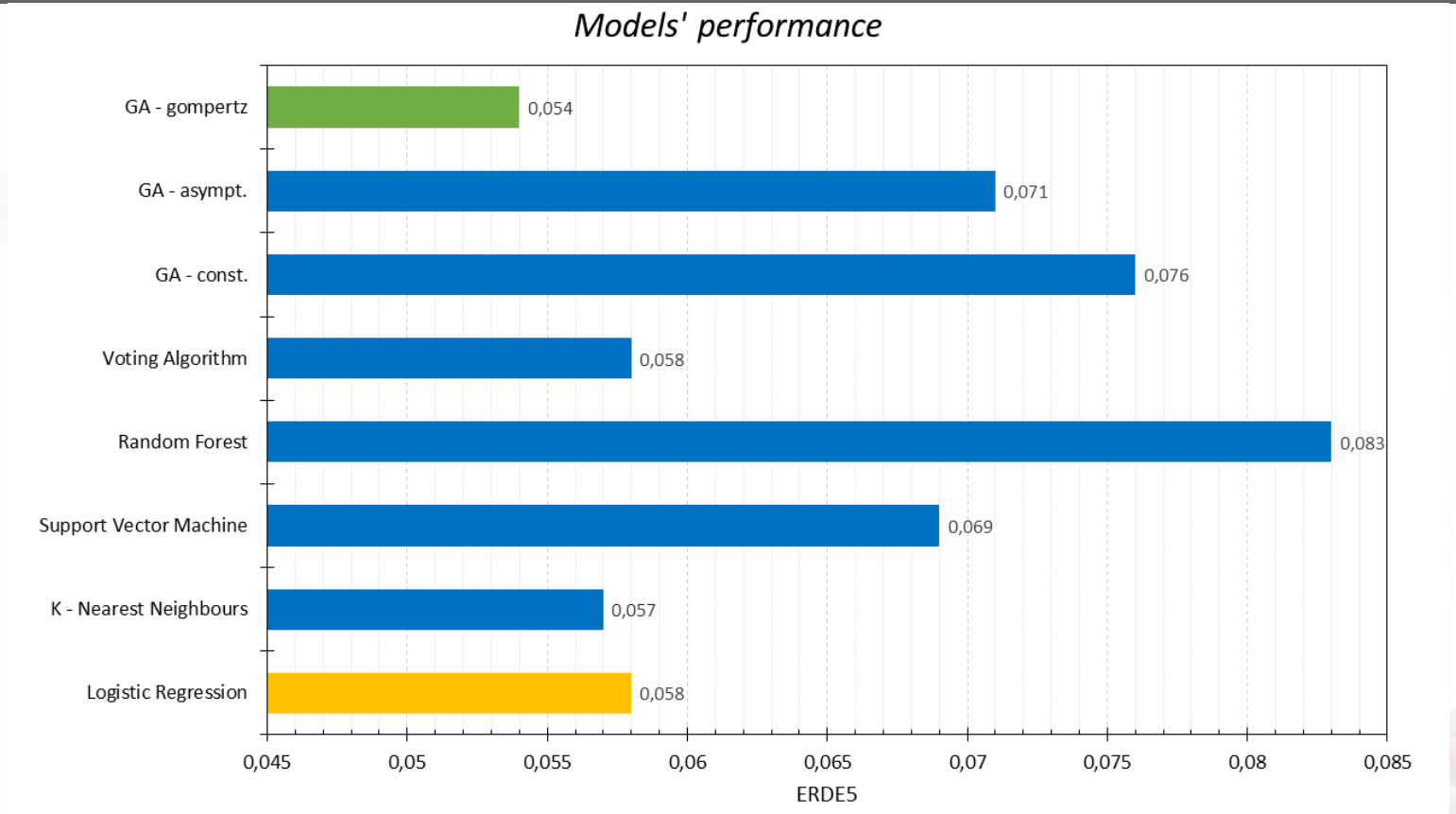
Mutation Rate functions:



		ERDE5	ERDE50	P	R	F1
B	Logistic Regression	0.058	**0.052**	0.41	0.77	0.54
LA	Voting Algorithm	0.058	0.053	0.50	0.69	0.60
LA + PCA	K - Nearest Neighbours	**0.057**	**0.052**	0.37	**0.79**	0.51
	Voting Algorithm	0.062	0.056	**0.58**	0.65	**0.61**
L+P+ V	Voting Algorithm	0.061	0.056	0.53	0.67	0.59
GA+ PCA+VADE	Constant	0.076	0.071	0.54	0.51	0.52
	Asymptotic	0.071	0.067	0.52	0.58	0.55
	Gompertz	0.069	0.063	0.53	0.60	0.56

		ERDE5	ERDE50	P	R	F1
B	Logistic Regression	0.058	0.052	0.41	0.77	0.54
LA	Voting Algorithm	0.058	0.053	0.50	0.69	0.60
LA + PCA	K - Nearest Neighbours	0.057	0.052	0.37	**0.79**	0.51
	Voting Algorithm	0.062	0.056	**0.58**	0.65	**0.61**
L+P+V	Voting Algorithm	0.061	0.056	0.53	0.67	0.59
GA+ PCA+VADE	Constant	0.076	0.071	0.54	0.51	0.52
	Asymptotic	0.071	0.067	0.52	0.58	0.55
	Gompertz	0.069	0.063	0.53	0.60	0.56
GA	Gompertz + PCA	0.061	0.055	0.53	0.67	0.59
	Gompertz + VADER	**0.054**	**0.048**	0.45	0.77	0.57

Summary


Models' performance
GA - gompertz
0,054
GA - asympt.
0,071
GA - const.
0,076
Voting Algorithm
0,058
Random Forest
0,083
Support Vector Machine
0,069
K - Nearest Neighbours
0,057
Logistic Regression
0,058
0,045
0,05
0,055
0,06
0,065
0,07
0,075
0,08
0,085
ERDE5

Conclusions

- This paper investigates how to better detect early risk of depression on social media, by optimizing time-aware classification measures: ERDE5 and ERDE50.
- We have applied different learning algorithms, and combinations of them.
- Other techniques such as dimensionality reduction and text polarity have been studied.
- We have provided further evidence for the benefit of applying genetic algorithms and text polarity (16.7% improvement regarding the baseline).

Future work

- Applying more accurate text sentiment classification to get a better representation of the input features.
- Exploring other pre-processing methods in order to reduce dimensionality.
- Trying the same approach with other datasets.

References

(Losada,2016) David E. Losada and Fabio Crestani. A test collection for research on depression and language use. In Lecture Notes in Computer Science, pages 28–39. Springer International Publishing, 2016.

(Li,2015) Ang Li, Xiaoxiao Huang, Bibo Hao, Bridianne O'Dea, Helen Christensen, and Tingshao Zhu. Attitudes towards suicide attempts broadcast on social media: an exploratory study of Chinese microblogs. PeerJ, 3:e1209, 2015.

(Park,2012) Minsu Park, Chiyoung Cha, and Meeyoung Cha. Depressive moods of users portrayed in twitter. In Proceedings of the ACM SIGKDD Workshop on Healthcare Informatics, pages 1–8, 2012.

(Cavazos-Rehg,2016) Cavazos-Rehg PA, Krauss MJ, Sowles S, Connolly S, Rosas C, Bharadwaj M, Bierut LJ. A content analysis of depression-related tweets. Comput Human Behav. 1;54:351-357, 2016.

(De Choudhury, 2013) De Choudhury, M., Counts, S., Horvitz, E.: Social media as a measurement tool of depression in populations. In: Proceedings of the 5th Annual ACM Web Science Conference, pp. 47–56. ACM (2013)

Towards Suicide Prevention:

Early detection of depression on social media

Víctor Leiva & Ana Freire

upf. Universitat Pompeu Fabra Barcelona

Unitat de Coordinació Acadèmica d'Enginyeries i Tecnologies de la Informació i les Comunicacions

EXCELENCIA MARÍA DE MAEZTU

ana.freire@upf.edu

@ana__freire