

Università degli Studi di Milano - Bicocca

Dipartimento di Informatica, Sistemistica e Comunicazione

Corso di laurea in Teoria e Tecnologia della Comunicazione

# Automatic detection of sexism on the web: visual analysis and multimodal classification

Relazione della prova finale di Monica Mantovani - 820453

Relatore Francesca Gasparini

Correlatore Elisabetta Fersini

Anno Accademico 2017-2018

# INTRODUZIONE

### Topic della tesi:

- Hate Speech
  - Sessismo
    - Meme

### Stato dell'arte:

Pubblicità sessiste

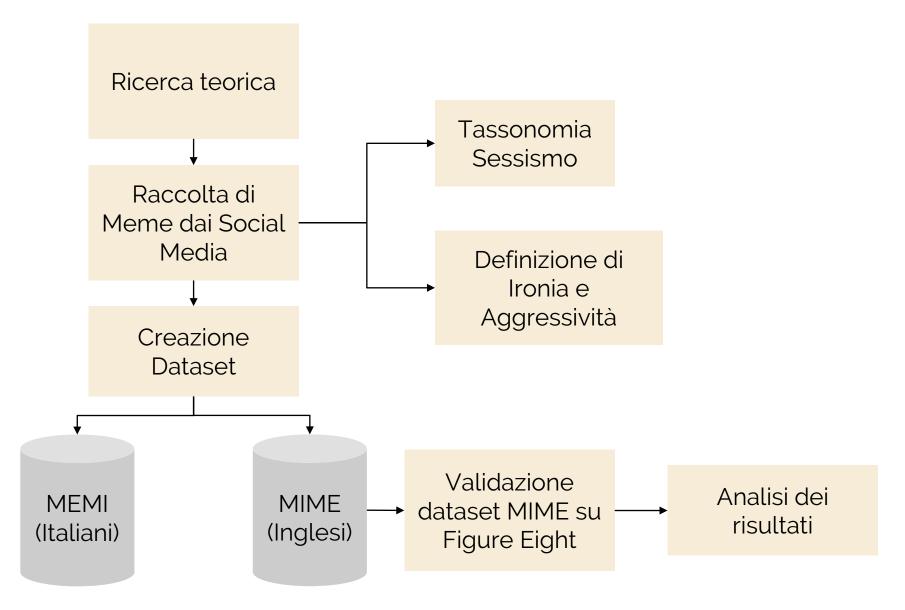
### Fasi del lavoro:

- Dataset
- Classificatore



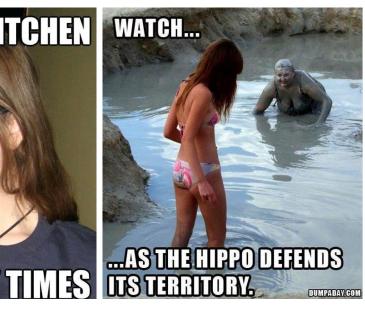


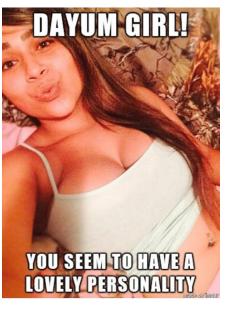
# PRIMA FASE: il dataset



# TASSONOMIA SESSISMO









Violenza

Shaming

Oggettificazione

Stereotipo

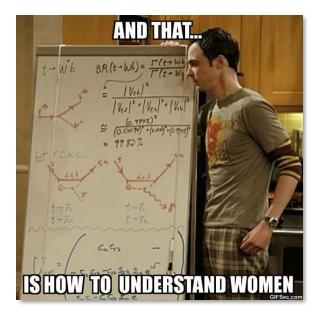
Inoltre, sono stati approfonditi aspetti di:

- Ironia -> definizione di meme
- Aggressività

# **DATASET**

800

**MIME** 



89

**Fonti** 

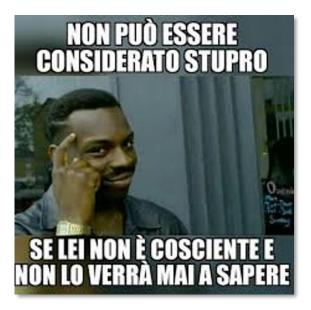
Blog

Forum

Social Network
Siti di meme

600

**MEMI** 



IT'LL BE FUN THEY SAID

# FIGURE EIGHT\*

60

Annotatori

In your opinion, is this meme sexist? (required)

- Yes
- O No

In your opinion, which part of the meme is sexist? (required)

- O Image
- O Text
- O Both

In your opinion, is this meme ironic? (required)

- O Yes
- O No

In your opinion, is this meme aggressive? (required)

- O Yes
- O No

GENERE:

- 52% maschi
- 48% femmine

ETA:

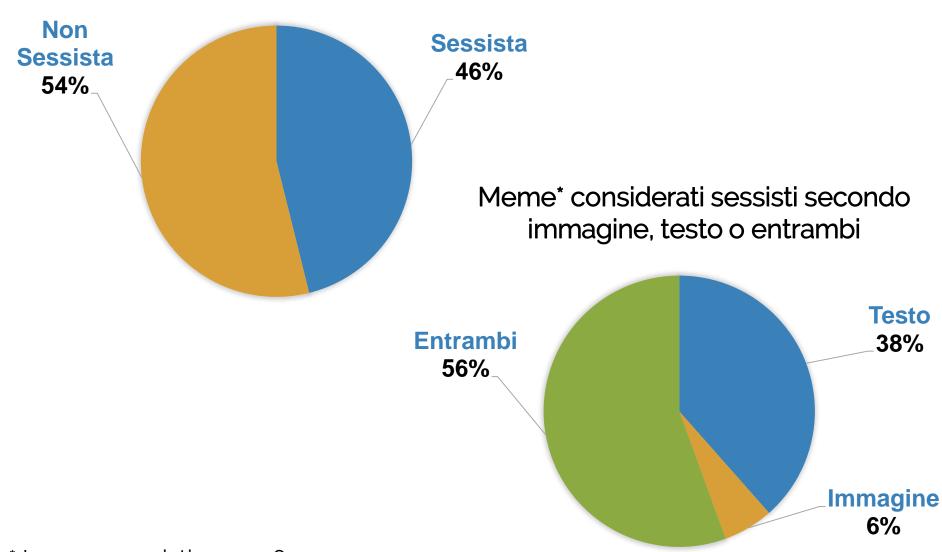
- 38%, 21-30
- 35%, 31-40
- 27%, 41-50

3 Annotatori per meme

<sup>\*</sup> Figure Eight: https://www.figure-eight.com/

# **ANALISI RISULTATI**

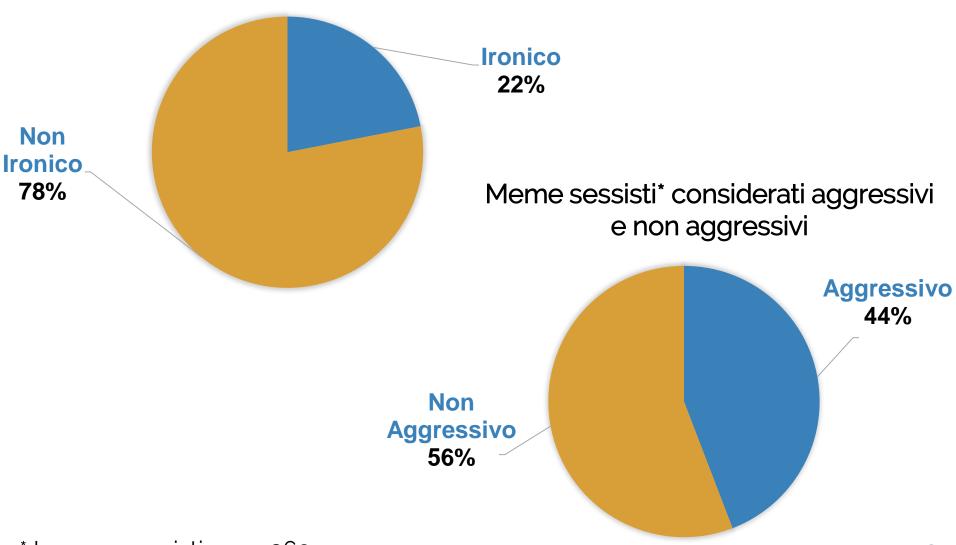
Meme considerati sessisti e non sessisti su 800



<sup>\*</sup> I meme sessisti sono 369

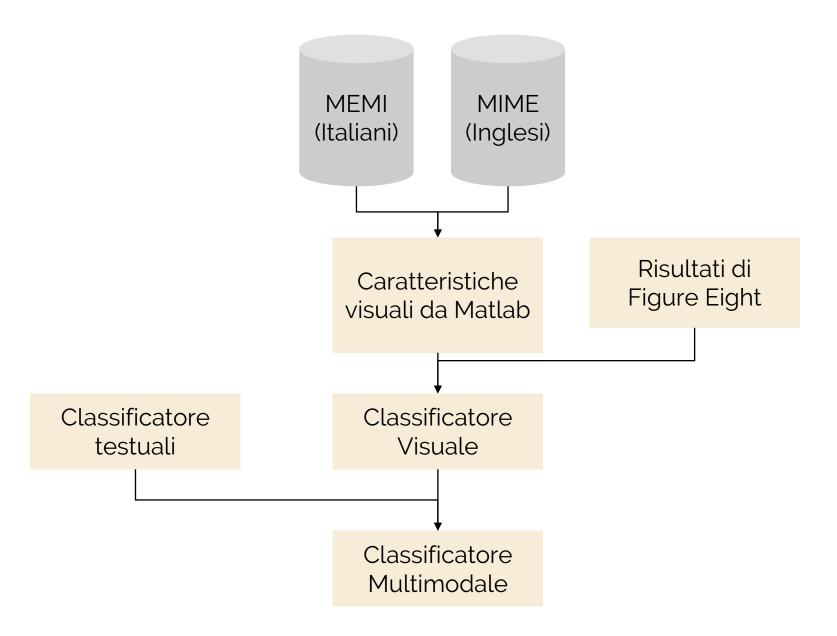
# **ANALISI RISULTATI (1)**

Meme sessisti\* considerati ironici e non ironici

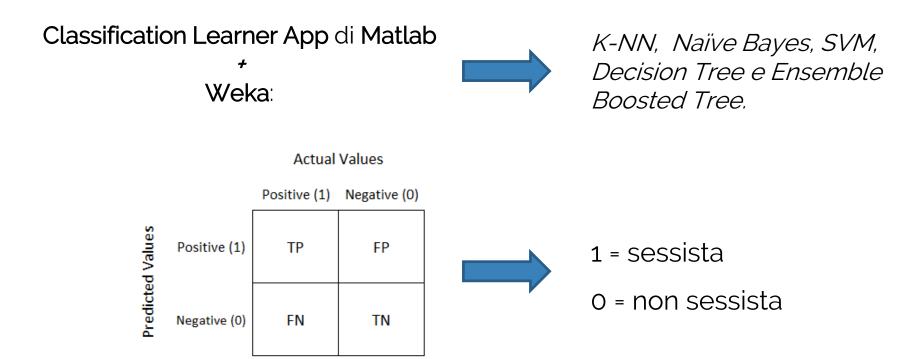


<sup>\*</sup> I meme sessisti sono 369

# SECONDA FASE: il classificatore



# CLASSIFICATORI



**Accuracy**: Percentuale di classificazioni corrette.

**Precision**: Percentuale di positivi trovati.

Recall: Percentuale di classificazione corretta nella classe corretta.

# **CLASSIFICATORE VISUALE**

# Training da SADit\* e test su MIME e MEMI:

- ▶ Hand-crafted
- ▷ Deep

Cross validation, 5 fold

Matlab

### Training e test su MIME:

- > Hand-crafted
- ▷ Deep

Cross validation, 5 fold

Matlab e Weka:

- Caratteristiche visuali
- Label

# LE LABEL









# Meme considerati sessisti secondo immagine, testo o entrambi Testo 38% Immagin e 6%

### Classificatore visuale:

- 1 = meme sessisti nelle
   immagini (22) e entrambi (142)
- 0 = meme sessisti nel testo (205)

### Classificatore multimodale:

- 1 = meme sessisti (369)
- 0 = meme non sessisti (431)

# CARATTERISTICHE VISIVE

### Caratteristiche hand-crafted:

### **Texture**

Coarseness, contrast, directionality, linelikness, roughness, edge density, LBP HoG

### Chromatic Properties

Chroma
Variance,
Number of
Regions,
Colorfulness,
Color
Histogram

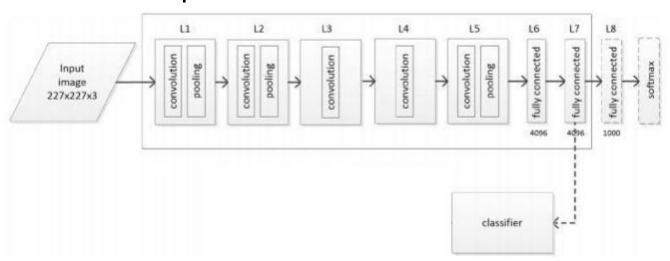
# Photographic Properties

features congestion, suuband entropy, image complexity, image blur, entropy, edge contrast, measure of enhancement

### Semantic Concepts

percentage of skin measure, number of faces.

### Caratteristiche deep (Metodo AlexNet\*):



\*Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks.

# PERFORMANCE: classificatore visuale

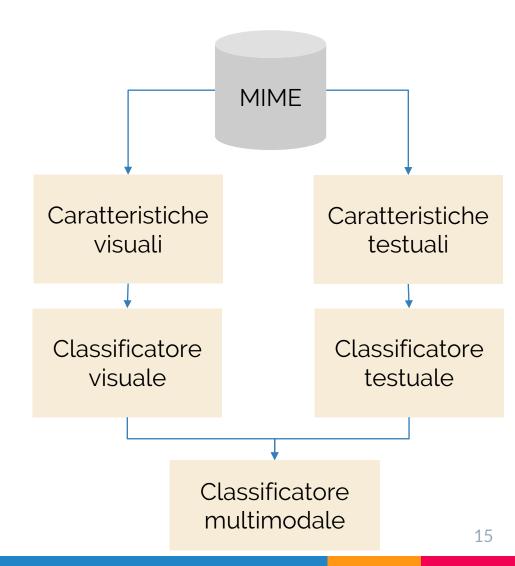
		Accuracy	Accuracy sessista	Accuracy non sessista	
Hand-Crafted	SADit su MIME	31%	87%	15%	
	SADit su MIME e MEMI	26%	84%	17%	
Deep	SADit su MIME	38%	78%	27%	
	SADit su MIME e MEMI	37%	78%	30%	
Weka	K-NN	75%	41%	84%	
	SVM	74%	36%	84%	
	Naïve Bayes	67%	47%	72%	
	Decision Tree	79%	7%	98%	
Matlab	SVM	80%	7%	99%	
	Ensemble Boosted Tree	66%	58%	68%	
	E.B.T. 70% + 30%	58%	52%	58%	

# CLASSIFICATORE MULTIMODALE

# **EARLY FUSION** MIME Caratteristiche Caratteristiche visuali testuali Classificatore

multimodale

### LATE FUSION



# PERFORMANCE di tutti i classificatori

		Immagine	Testo	Early Fusion	Late Fusion
	Accuracy	67%			71%
Naïve Bayes	Recall	47%			82%
	Precision	84%			64%
K-NN	Accuracy	75%		52%	74%
	Recall	41%		49%	79%
	Precision	84%		56%	70%
SVM	Accuracy	74%	76%	68%	75%
	Recall	36%	73%	65%	79%
	Precision	83%	79%	70%	70%
Decision Tree	Accuracy	79%			77%
	Recall	7%			72%
	Precision	80%			77%

# IL MIO LAVORO

- ✓ Tassonomia Sessismo
- ✓ Dataset: MIME e MEMI
- ✓ Esperimento per validazione MIME su Figure Eight
- ✓ Analisi risultati
- ✓ Caratteristiche visuali su Matlab
- ✓ Classificatore visuale su Matlab
- ✓ Classificatore visuale su Weka
- ✓ Classificatore multimodale, approccio Early e Late Fusion su Weka

# CONCLUSIONI

### PRIMA FASE:

- Meme: Immagini uguali con testo diverso
- Esperimento:
- Investigare aspetto ironia e aggressività
- Numero maggiore di meme
- Numero maggiore di utenti
- Differenza generazionale
- Differenza di genere
- Livello di istruzione: eventi storici o attuali

### **SECONDA FASE:**

- Problema riconoscere immagini sessiste -> contesto
- Immagini diverse; sintetiche o reali
- Caratteristiche visuali da migliorare
- Migliori risultati con Late fusion
- Migliorare approccio Early

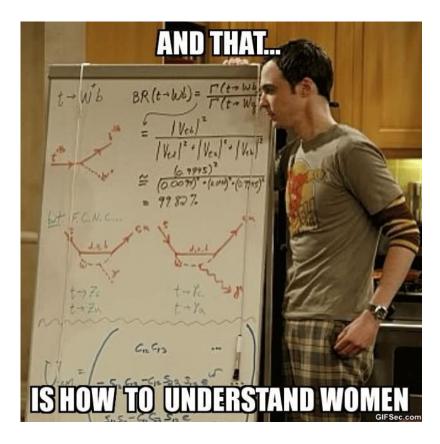
# Grazie per l'attenzione!

# **IRONIA**

### «Affermare l'esatto contrario di ciò che si vuol dire»

- -> ironia situazionale
- -> ironia pittorica

'Pubblicità e ironica se un elemento, visuale o testuale, è incosistente con l'altro, scioccando l'audience.'



Lombardi, M. (2018). *Il nuovo manuale di tecniche pubblicitarie*. COMMunity - Corporate & Marketing Communication.

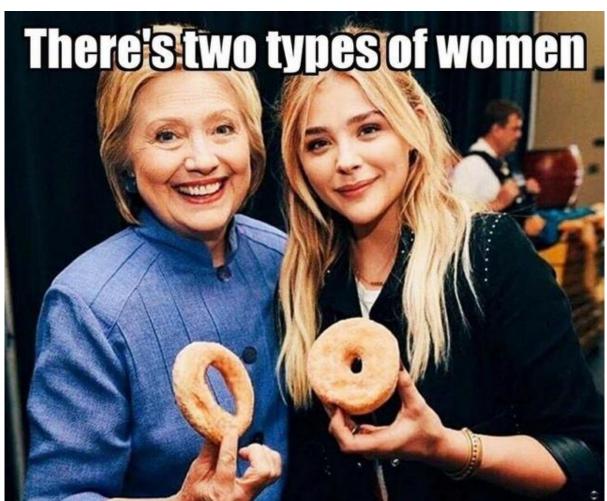
# **SESSISMO**

Una forma di discriminazione nei confronti delle donne.

-> misoginia internalizzata

### Meme sessista:

- Immagini
- Testo
- Entrambi



# SESSISMO ED ESPERIMENTO

Scala Likert (6 valori)



Quanto trovi sessista il meme? E' abbastanza sessista? E' troppo sessista?

VALIDAZIONE DATASET

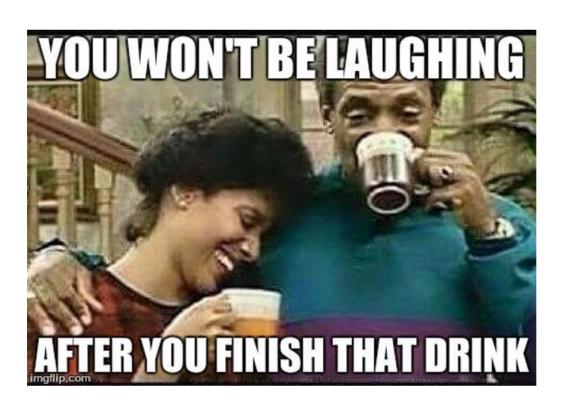
# IRONIA ED ESPERIMENTO

Ironia: il meme è ironico ma non fa ridere

«In your opinion, is this meme ironic?(required)»

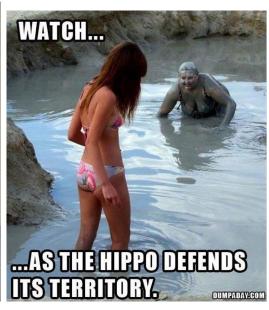
- > Lingua
- Conoscenza attualità
- > Livello culturale
- > Età

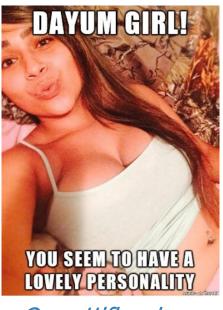
VALIDAZIONE DATASET



# TASSONOMIA SESSISMO









Violenza

Shaming

Oggettificazione

Stereotipo

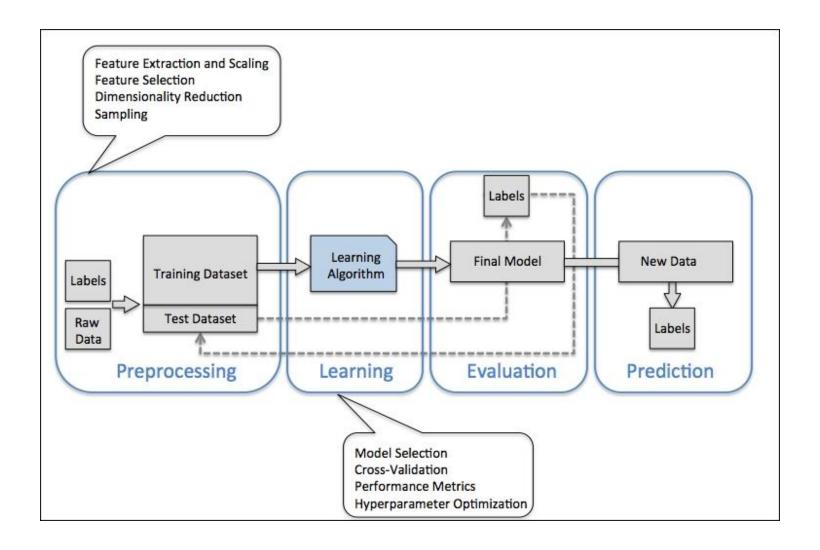
Body shaming Slut shaming

Di genere Di ruolo

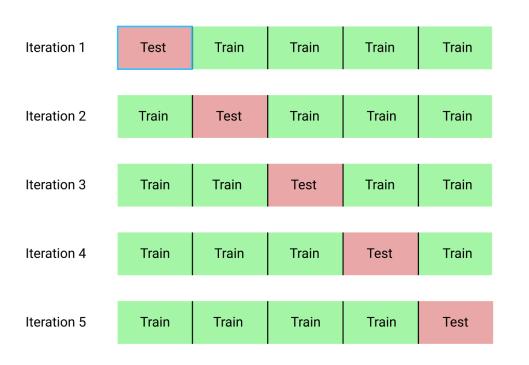
Rape culture

Oggettificazione sessuale Self-objectification

# CLASSIFICATORE



## **CROSS-VALIDATION**



La convalida incrociata (*cross-validation* in inglese) è una tecnica statistica utilizzabile in presenza di una buona numerosità del campione osservato o *training set*.

-> k-fold cross-validation consiste nella suddivisione del dataset totale in k parti di uguale, si esclude iterativamente un gruppo alla volta e lo si cerca di predire con i gruppi non esclusi. Ciò al fine di verificare la bontà del modello di predizione utilizzato.

# LATE FUSION

ppunti	isi C	Acie 18	A	IIIIeailieillo		(3) (	Numen	(3) (			31	ш
<b>*</b> :	$\times$ $\checkmark$ $f_x$	=SE(F2=1;1;SE(J2=1	;1;0))									
Е	F	G	н	1	J	K	L	М	N	0	Р	
att1	prediction testo	imma <sub>b</sub>	ımmagini svm i	mmagini bayes	immagini kr	nn	testo+tree	testo+svm	testo+bay	testo+knn	ground_truth_all	
ENG01	1	. 0	0	1	0		1	1	1	1		1
ENG02	1	. 0	1	0	0		1	1	1	1		1
ENG03	1	. 0	0	0	0		1	1	1	1		1
ENG04	1	. 0	0	0	0		1	1	1	1		1
ENG05	1	. 0	1	1	1		1	1	1	1		1
ENG06	1	. 0	0	0	0		1	1	1	1		1
ENG07	1	. 1	1	1	1		1	1	1	1		1
ENG08	1	. 0	0	0	0		1	1	1	1		1
ENG09	0	1	1	1	1		1	1	1	1		1
ENG10	1	. 0	0	0	0		1	1	1	1		1
ENG100	1	. 0	0	0	0		1	1	1	1		1
ENG101	1	. 0	0	1	1		1	1	1	1		1
ENG102	1	. 0	0	0	0		1	1	1	1		1
ENG103	1	. 0	0	0	0		1	1	1	1		1
ENG104	1	. 0	0	0	0		1	1	1	1		1
ENG105	1	. 0	0	0	0		1	1	1	1		1
ENG106	1	. 0	0	0	0		1	1	1	1		1
ENG107	1	. 0	0	1	1		1	1	1	1		1

- Creato label con condizione SE
- Unite le label con caratteristiche visuali e testuali
- Creato classificatore