



Corso di laurea magistrale in Data Science per
l'economia e le imprese

Sviluppo di modelli per l'emotion recognition di testi in lingua italiana

Relatore:

Prof. Potena Domenico

Tesi di Laurea di:

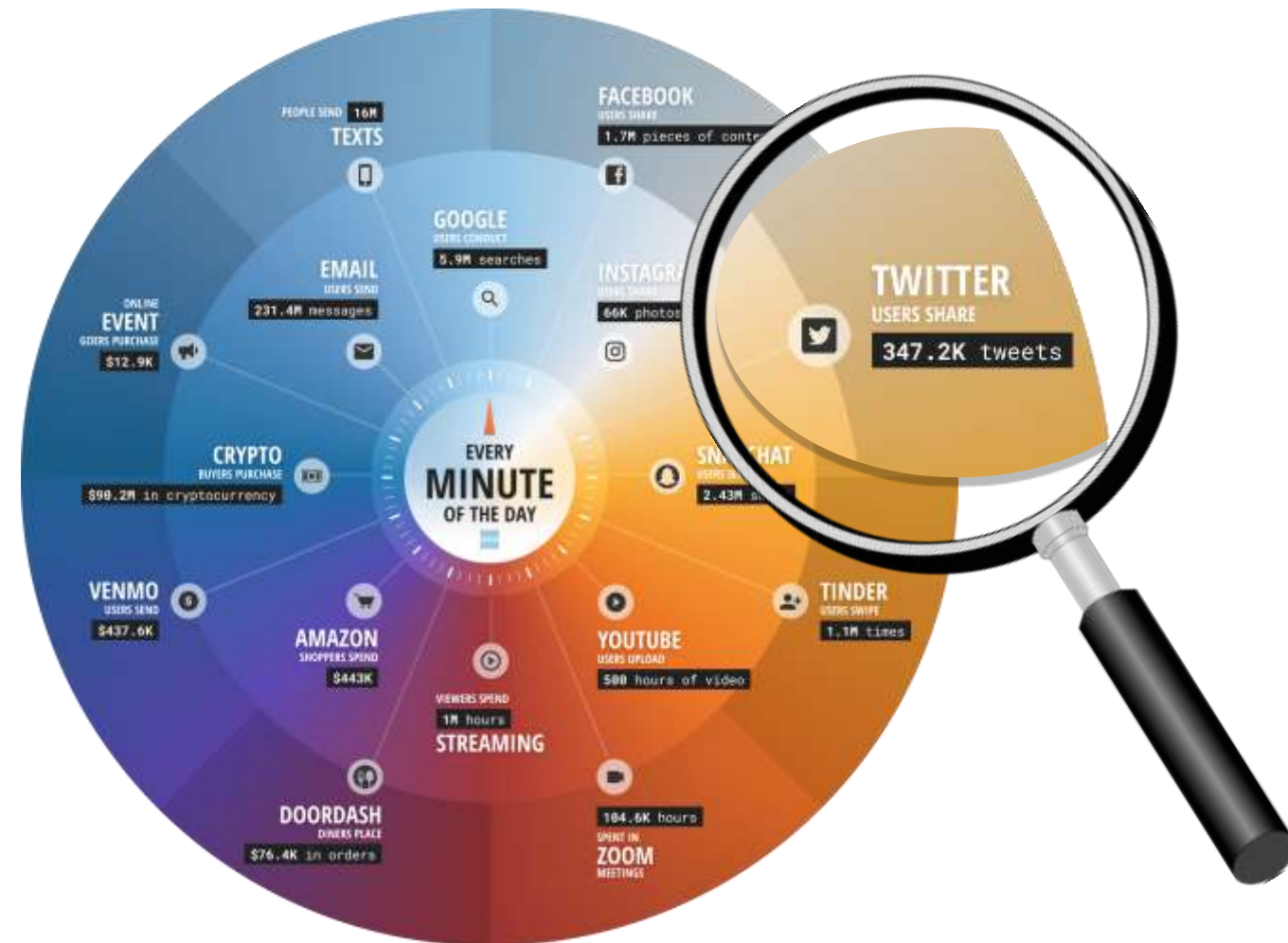
Ridolfi Lorenzo

Correlatore:

Prof. Mircoli Alex

Contesto di riferimento

- Social network
 - Twitter
- Big Data
 - Grandi volumi
 - Elevata velocità



Motivazioni

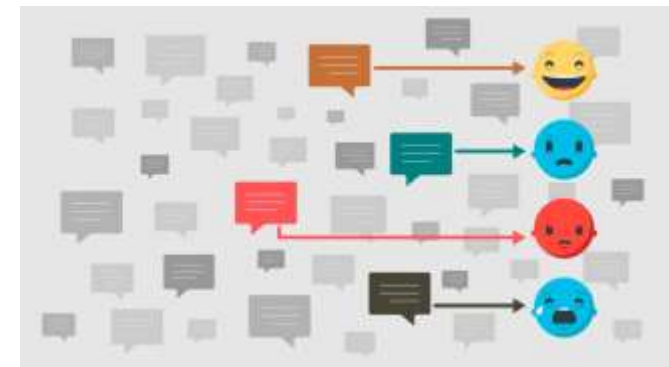


Entity Recognition

Automatically find names
of people, places, products,
and organizations in text
across many languages.



Emotion Recognition

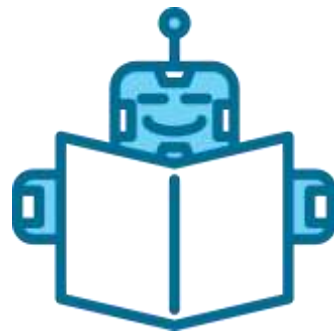


Obiettivo

- **Sviluppo di modelli per l'analisi automatica del testo**
- Intelligenza Artificiale
- Natural Language Processing



Metodologia



1. Preprocessing

2. Training

3. Validation

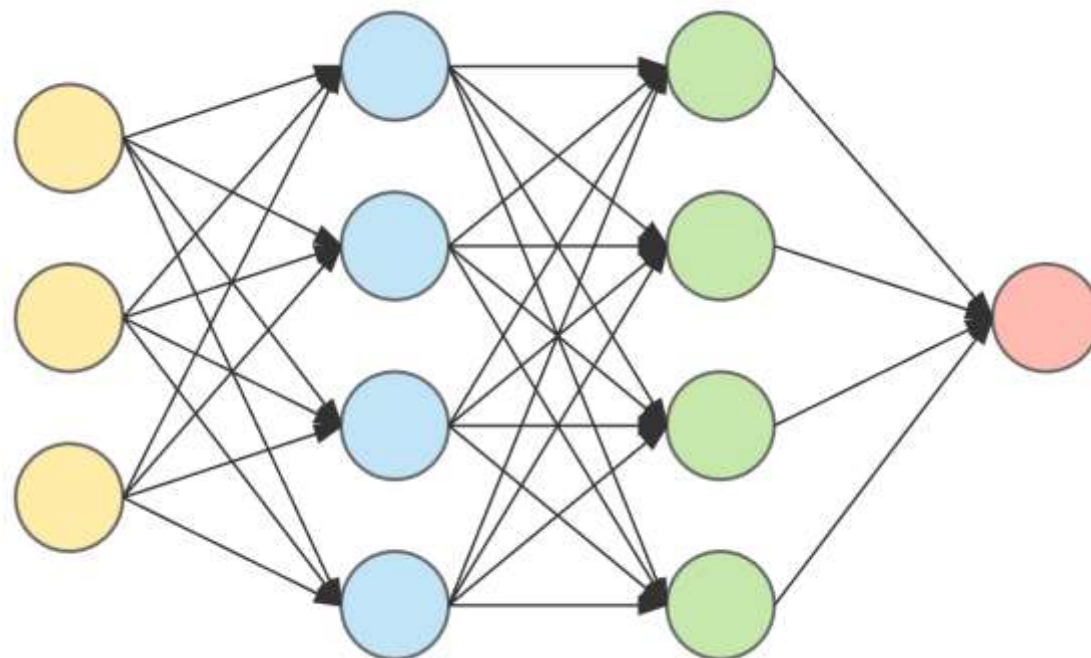
Training

GPT-3

- Fine Tuning
 - Ada
 - Davinci



- BERT multilingual
- Fully connected
 - Dense Layers
 - Learning rate
 - Max length
 - Epochs



Validation

- TwlT: Dataset con tweets in italiano

- 3108 records
- Classi:
 - 0 Happiness
 - 1 Trust
 - 2 Sadness
 - 3 Anger
 - 4 Fear
 - 5 Disgust

- Hold-out: 80-20



		Text	Emotion
Id			
2932	Hanno imparato bene dall'America. \r\nQuesta è...		5
2125	Io non ci voglio più venire a Roma		4
136	Il buongiorno che speravo di dare da molto tem...		0
2154	Subito Mary, non prendere mi raccomando		4
678	Già non mi perdevo una puntata prima figurati ...		1
1999	Ti odiamo proprio, altro che cattiva		3
610	Ciao, molto bello		1
561	GRANDE LISO \r\nGIU' LE MANI DALLE NOSTRE TERR...		1
2077	ah, ad esempio io ancora non sapevo che l'Asal...		4
399	Finalmente ho iliad		0

Risultati raggiunti

GPT-3

	Support	Precision	Recall	F1-Score
Happiness	117	0.94	0.96	0.95
Trust	90	0.97	0.93	0.95
Sadness	92	0.86	0.90	0.88
Anger	104	0.86	0.88	0.87
Fear	114	0.92	0.82	0.87
Disgust	101	0.97	0.93	0.90
Accuracy				0.90

The logo for Google's BERT model, featuring a stylized 'G' in Google's signature colors (red, yellow, green, blue) above the word 'BERT' in red capital letters.

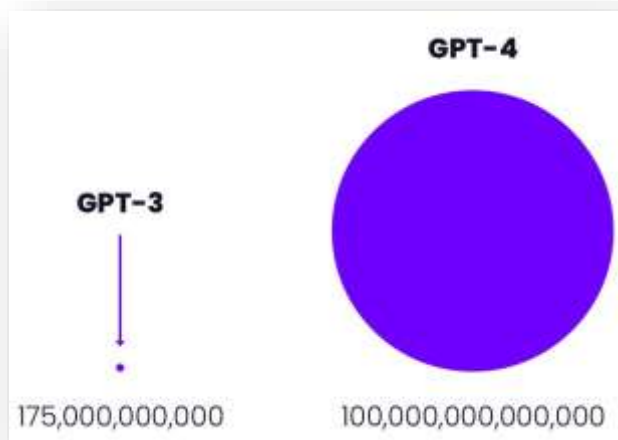
	Support	Precision	Recall	F1-Score
Happiness	117	1.00	0.93	0.96
Trust	90	0.93	0.97	0.95
Sadness	92	0.92	0.86	0.89
Anger	104	0.75	0.82	0.82
Fear	114	0.92	0.85	0.85
Disgust	101	0.86	0.89	0.89
Accuracy				0.89

Conclusioni

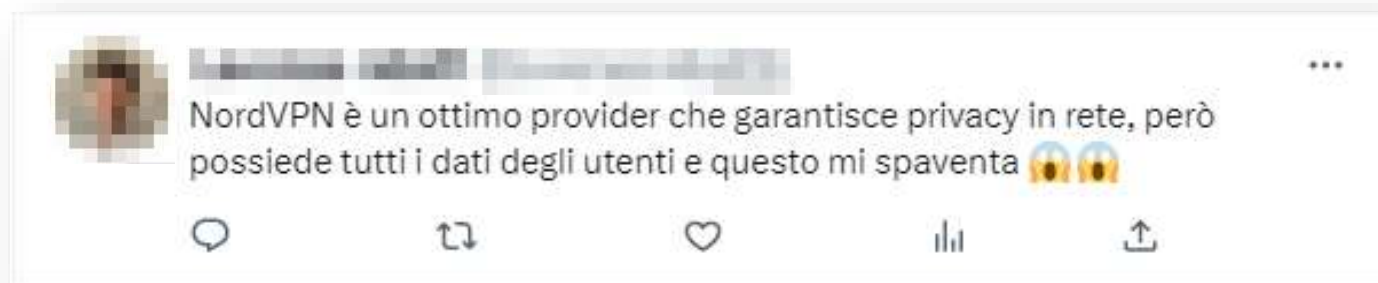
- Modelli di classificazione accurati

- Sviluppi futuri

- GPT-4



- Tweet multi-emotions



Grazie per l'attenzione!



<https://github.com/lorenzoridolfi9/Emotion-Recognition>