crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Sentiment Analysis con il CREA: A Deep Learning Approach

SPEAKER
Ph.D. Marco Polignano
Assistant Professor
Università degli Studi di Bari Aldo Moro
70017, Bari, Italia

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Natural Language

Refers to the **language spoken by people**, e.g. English, Japanese, Swahili, Italian, as opposed to artificial languages, like C++, Java, etc.

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Processing ...

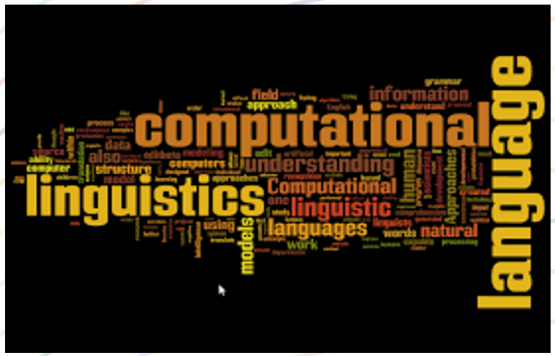
Applications that deal with natural language in a way or another

- processing language with computers
- **go beyond the keyword matching**: identify the **structure** and **meaning** of words, sentences, texts and conversations

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

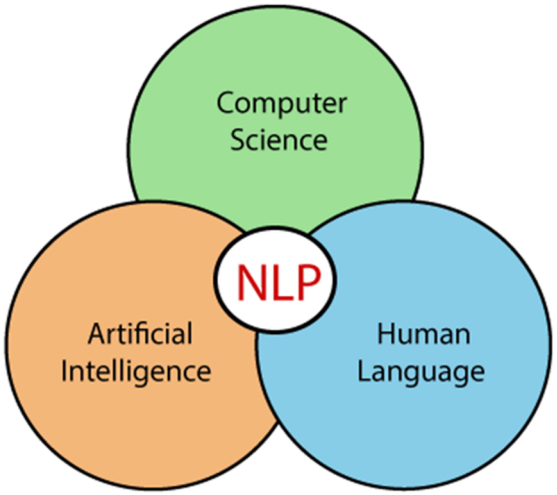
Computational Linguistics

- Scientific study of language from a computational perspective
  - modeling linguistic problems through computational procedures



crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

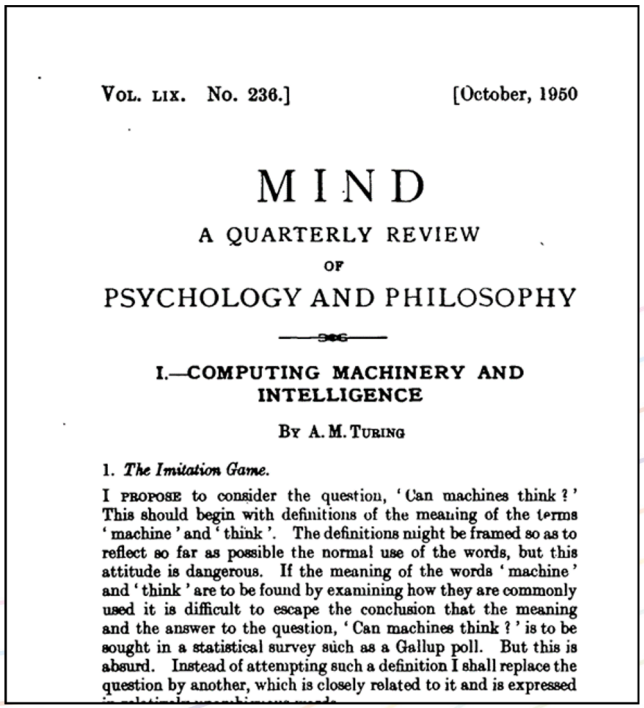
What is Natural Language Processing ?



Field of **computer science**, **artificial intelligence**, and **computational linguistics** concerned with the interactions between computers and human languages

crea
Consiglio per la ricerca in agricoltura
e l'analisi dell'economia agraria

NLP is an AI-complete task



Vol. lix. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. *The Imitation Game.*

I propose to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed

"Can machines think?"

Turing defined thinking with an operational method involving the **use of language**

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Where we are now …

Let's ask "what is NLP?" to an AI system!

ChatGPT is a chatbot by OpenAI that uses NLP techniques to maintain a conversation with a user in natural language.

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

M Cosa è l'NLP?

NLP, o Elaborazione del Linguaggio Naturale (dall'inglese Natural Language Processing), è un campo dell'intelligenza artificiale che si occupa di consentire alle macchine di comprendere, interpretare e generare linguaggio umano in modo simile a come lo farebbe un essere umano.

L'obiettivo principale dell'NLP è di creare sistemi in grado di interagire con gli esseri umani utilizzando il linguaggio naturale, permettendo alle macchine di comprendere testi, discorsi, conversazioni e produrre risposte appropriate in modo autonomo.

Le applicazioni dell'NLP sono ampie e includono:

1. **Chatbot e Assistenti Virtuali**: Creazione di programmi in grado di conversare con gli utenti in modo naturale, aiutandoli a ottenere informazioni o a svolgere compiti specifici.
2. **Traduzione Automatica**: Sistemi in grado di tradurre testi da una lingua all'altra.
3. **Analisi del Sentimento**: Valutazione automatica del tono emotivo di un testo, utile per comprendere l'opinione o il sentiment degli utenti.
4. **Estrazione di Informazioni**: Identificazione automatica di informazioni rilevanti all'interno di testi (ad esempio, estrazione di dati da curriculum, articoli, ecc.).

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

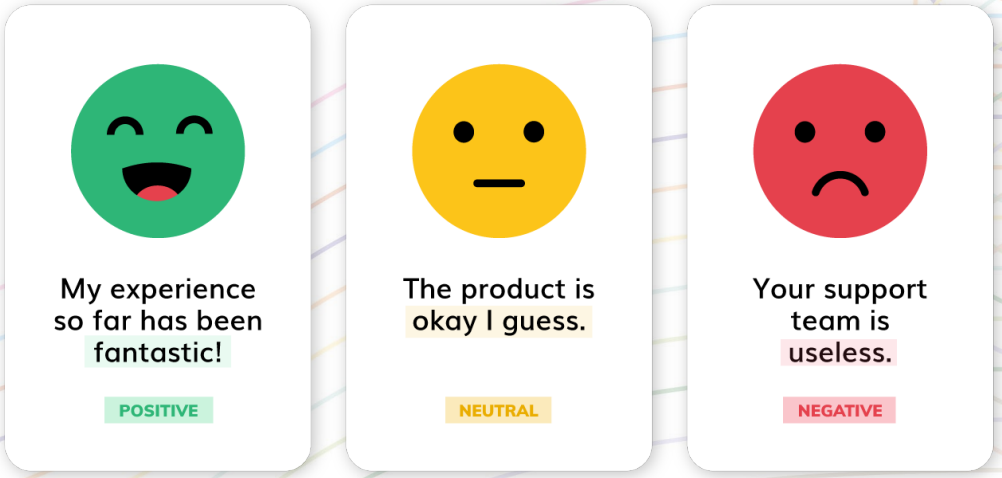
Why NLP?

Big Giants have strong NLP Labs!
Lots of commercial applications and interest!

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Sentiment Analysis

Also known as *opinion mining*, is the task of identifying the **subjectivity** (neutral vs. emotionally loaded) and the **polarity** (positive vs. negative semantic orientation) of a text, by exploiting natural language processing and computational linguistics.



crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Aspect-based Sentiment Analysis

- Identification of the aspects of given target entities and the sentiment expressed for each aspect
- Summarization of the content of users' reviews

*"The **food** was lousy - too sweet or too salty and the **portions** tiny"*

The sentiment with respect to these two aspects is negative

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Stance Detection

Automatically determining from text whether the author is in favor of the given target, against the given target, or whether neither inference is likely.

Target: legalization of abortion
Tweet: *A foetus has rights too! Make your voice heard.*

Humans can deduce from the tweet that the speaker is likely against the target.

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Stance Detection

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Knowledge Acquisition Bottleneck

- **Problem in Decision Support Systems (DDS)**
  - Experts not willing to elicit their knowledge
  - Experts unable to formalize their knowledge
  - Experts ignore details of their knowledge
  - People do not know why they take a decision
- Solution: **Machine Learning (ML)**
  - Learning
    - Knowledge Acquisition
  - Machine
    - Carried out by computers

Learning

Classical view

- A system
  - Natural or artificial

learns if, using its experience, changes something in its behavior/operation so as to *improve* its own *performance* in

- Solving a problem
- Attaining a goal
- Carrying out a task

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Machine Learning

Reproduction on a computer of the intelligent learning capabilities:

- Emulation: Strong AI
- **Simulation: Weak AI**

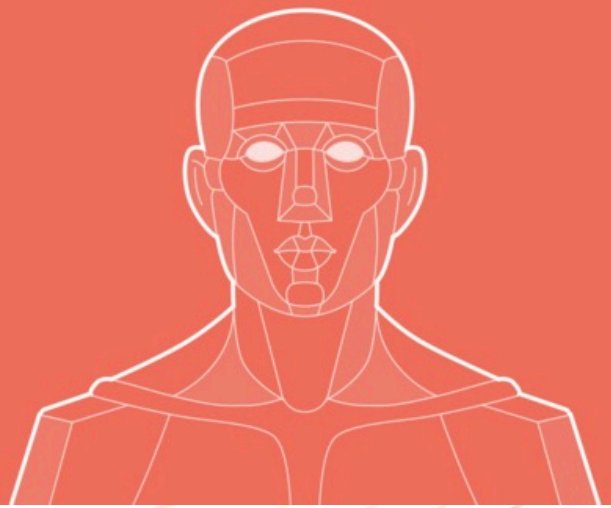
> *"That may be exactly what's needed for anybody who wants to go into this field [AI], namely, blind optimism with no reasonable basis for it."*

Raj Reddy (Turing prize 1994)

crea
Consiglio per la ricerca in agricoltura
e l'analisi dell'economia agraria

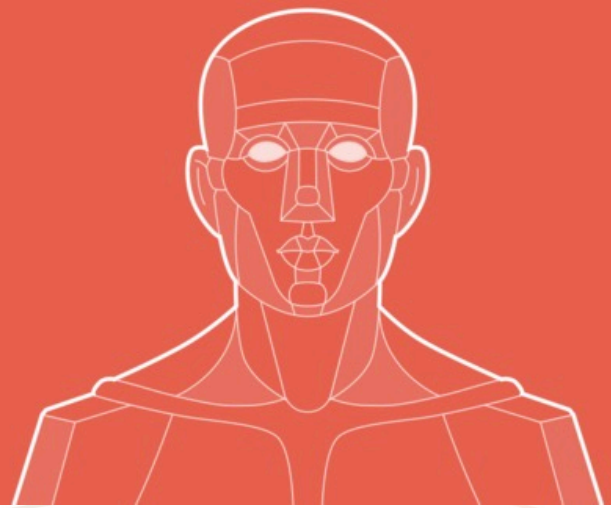
Simulation

IMITATE, MIMIC, PRETEND,
'GIVE THE APPEARANCE OF'



Emulation

REPRODUCE OR DUPLICATE THE FUNCTIONS OF A SYSTEM IN A WAY THAT IS FUNCTIONALLY IDENTICAL TO THE THING BEING EMULATED.



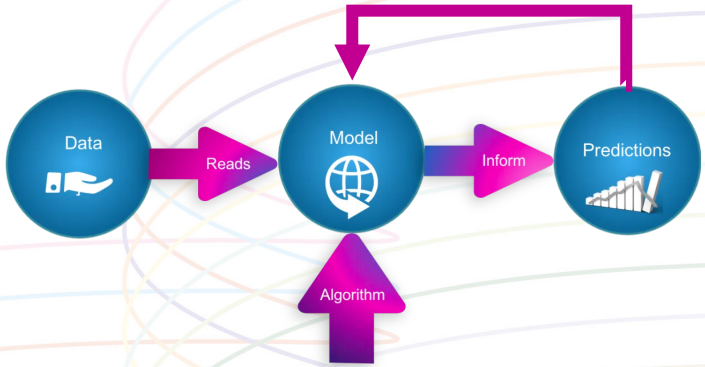
Machine Learning

A program

- Learns from **experience E**
- With respect to a given **set of tasks T**
- And with a **performance measure P**

if

- Its **performance on tasks T** Measured
- through P
- **Improves with experience E**



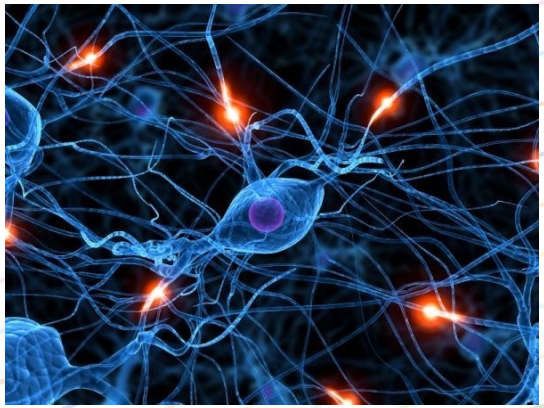
Any learning program must identify and define

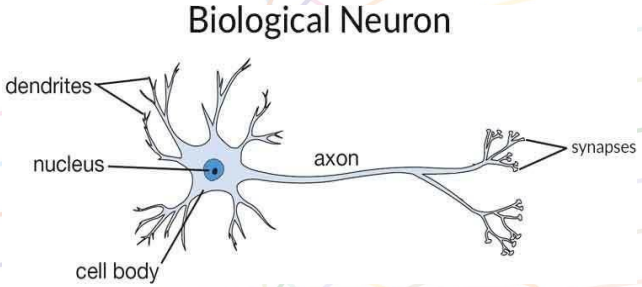
- The class of tasks
- The performance measure to improve
- The source of experience

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Artificial Neural Networks

- Inspired to the structure of the human brain
  - Neurons as elaboration units
  - Synapses as connection network

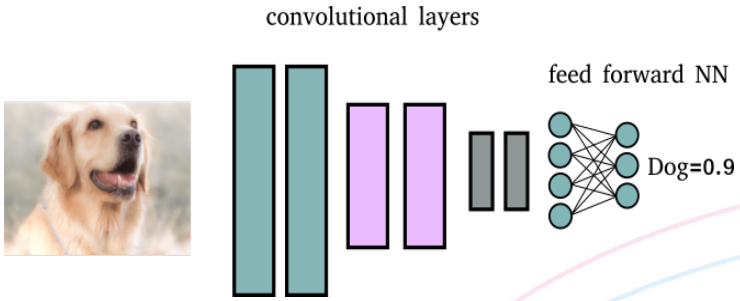




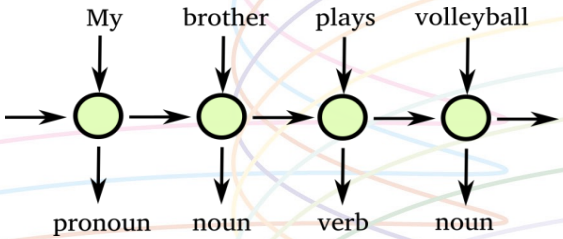
Artificial Neural Networks

- Different tasks, different architectures

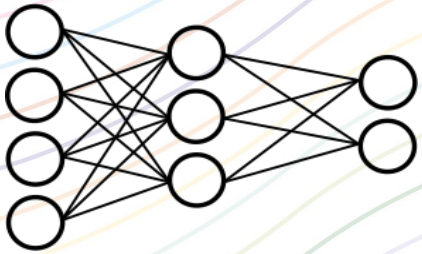
Image understanding: Convolutional NN (CNN)



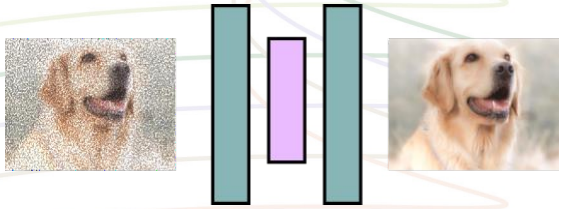
Time series analysis: Recurrent NN (RNN)



Numerical vectors classification: feed forward NN (FFNN)

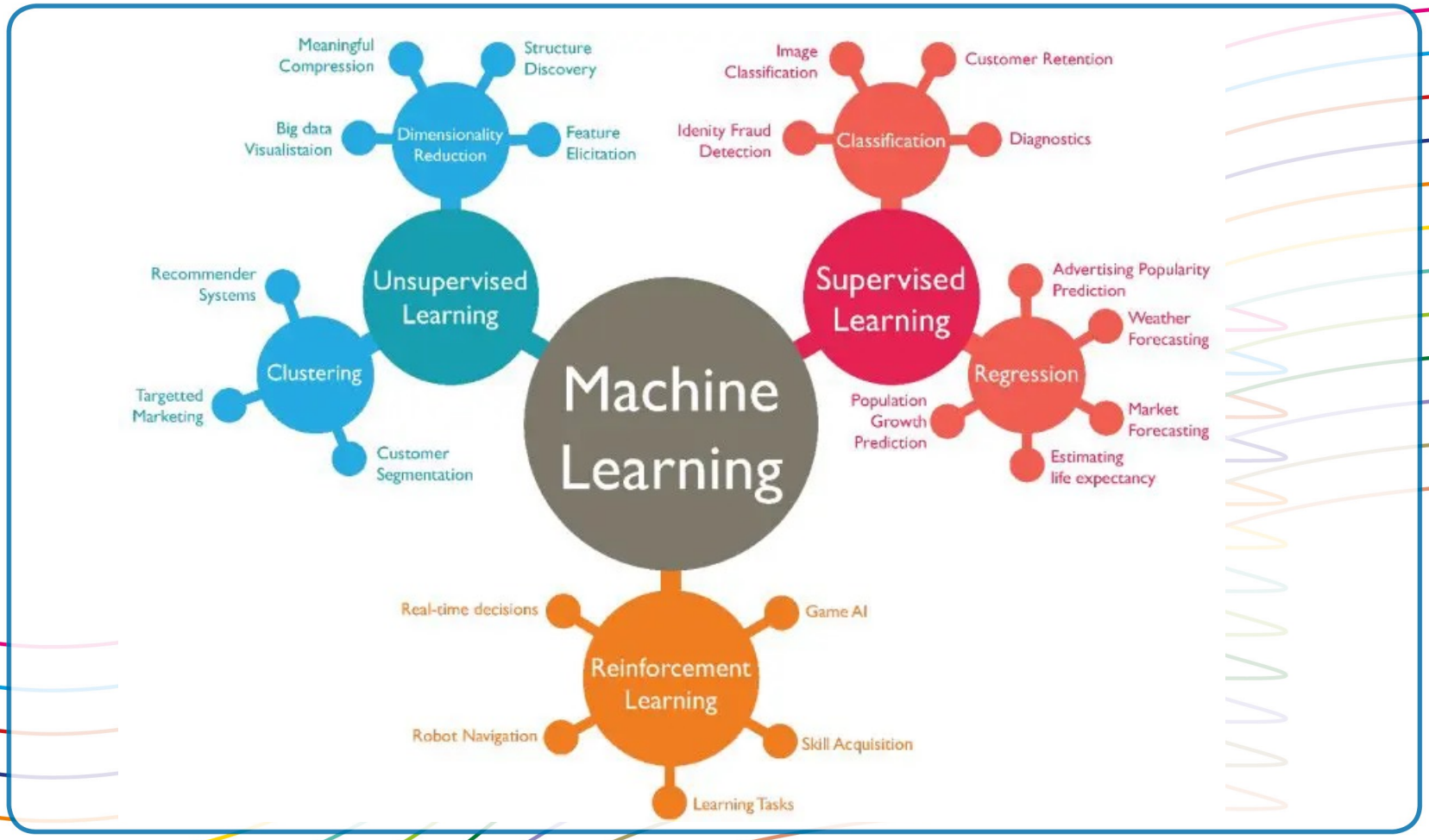


Denoising: auto-encoders



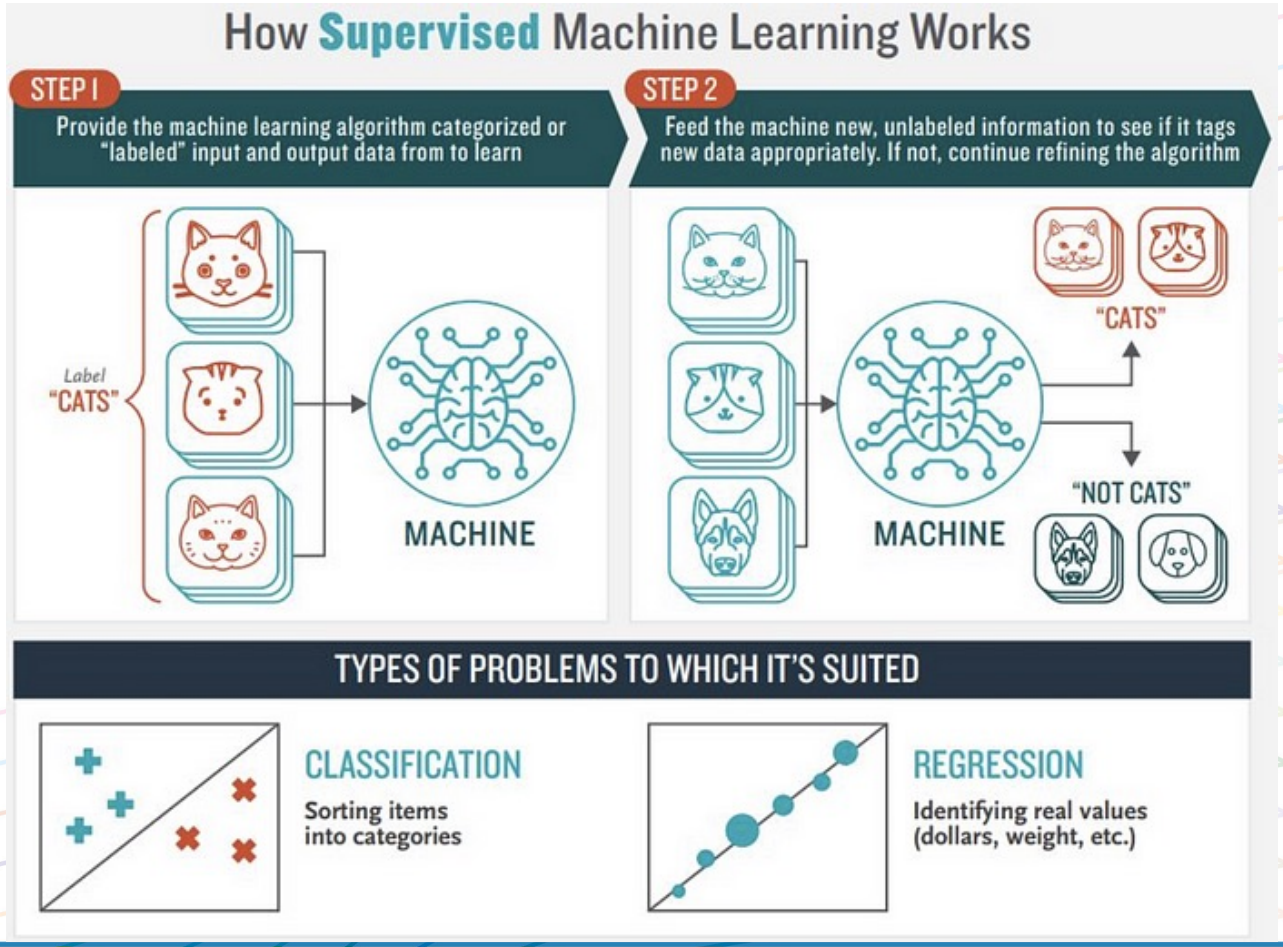
crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Machine Learning



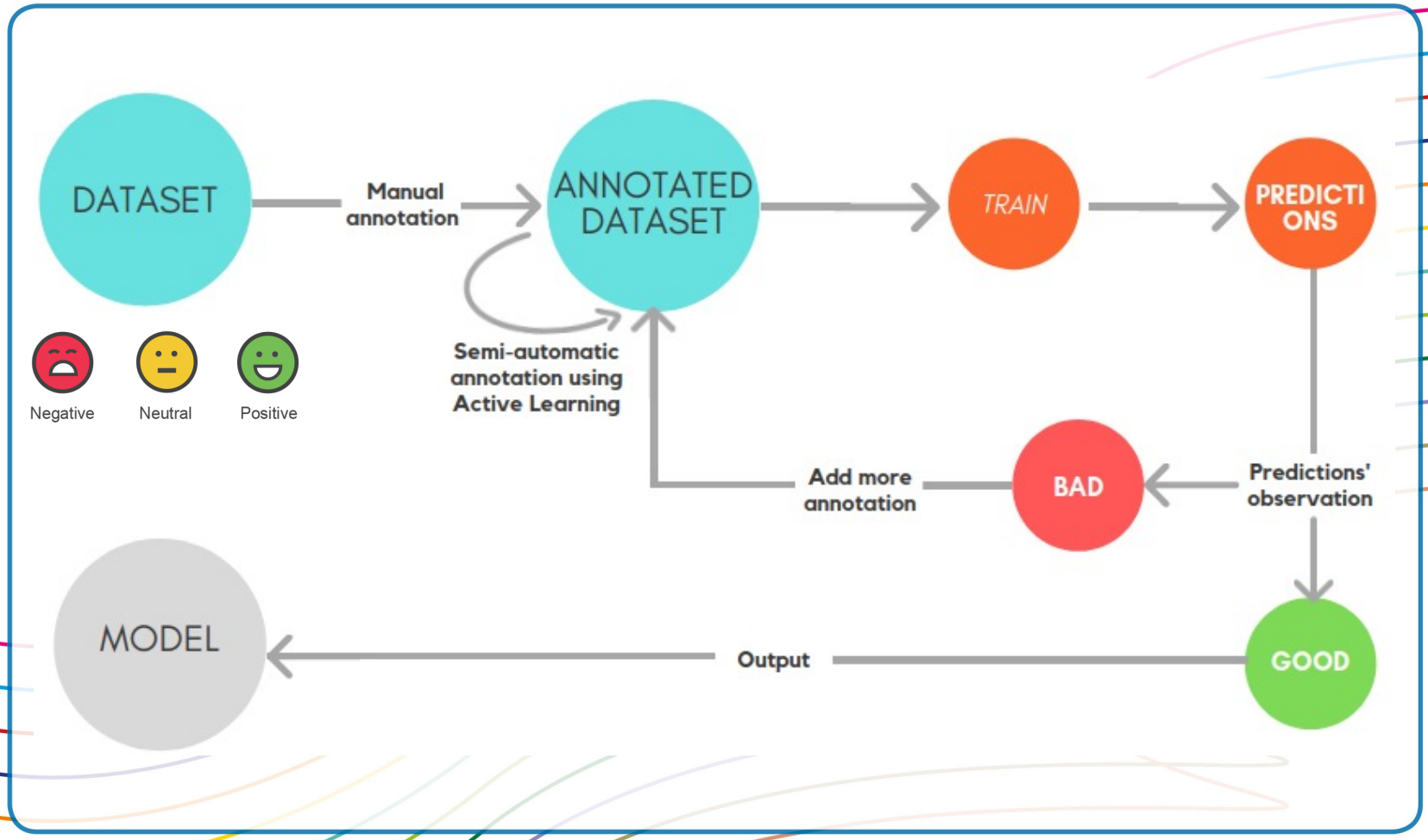
crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Supervised Machine Learning



crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

ML Pipeline



Vassallo M., Gabrieli G., Basile V., Bosco C. The tenuousness of lemmatization in lexicon-based sentiment analysis. Proceedings of the sixth Italian conference on computational linguistics, Bari, Italy, November 13-15, 2019, CEUR workshop proceedings, vol. 2481, CEUR-WS.org (2019)

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

AGRITrend

In total **AGRITrend** provides annotations for **1000 tweets about agriculture**.

AGRITrend is annotated with two labels *pos* and *neg.*

- If pos and neg are both 0, the tweet is considered as neutral.
- If both the categories are annotated as 1, the tweet is considered mixed

@SyngentaItalia Disponibile il catalogo di #Syngenta per il 2019. Si rafforza ulteriormente l'offerta di sementi, agrofarmaci e tecnologie per meglio rispondere alle sfide dell'agricoltura https://t.co/etBQba2vpu di @AgriculturaIT #ricerca #agricoltura #syngenta

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

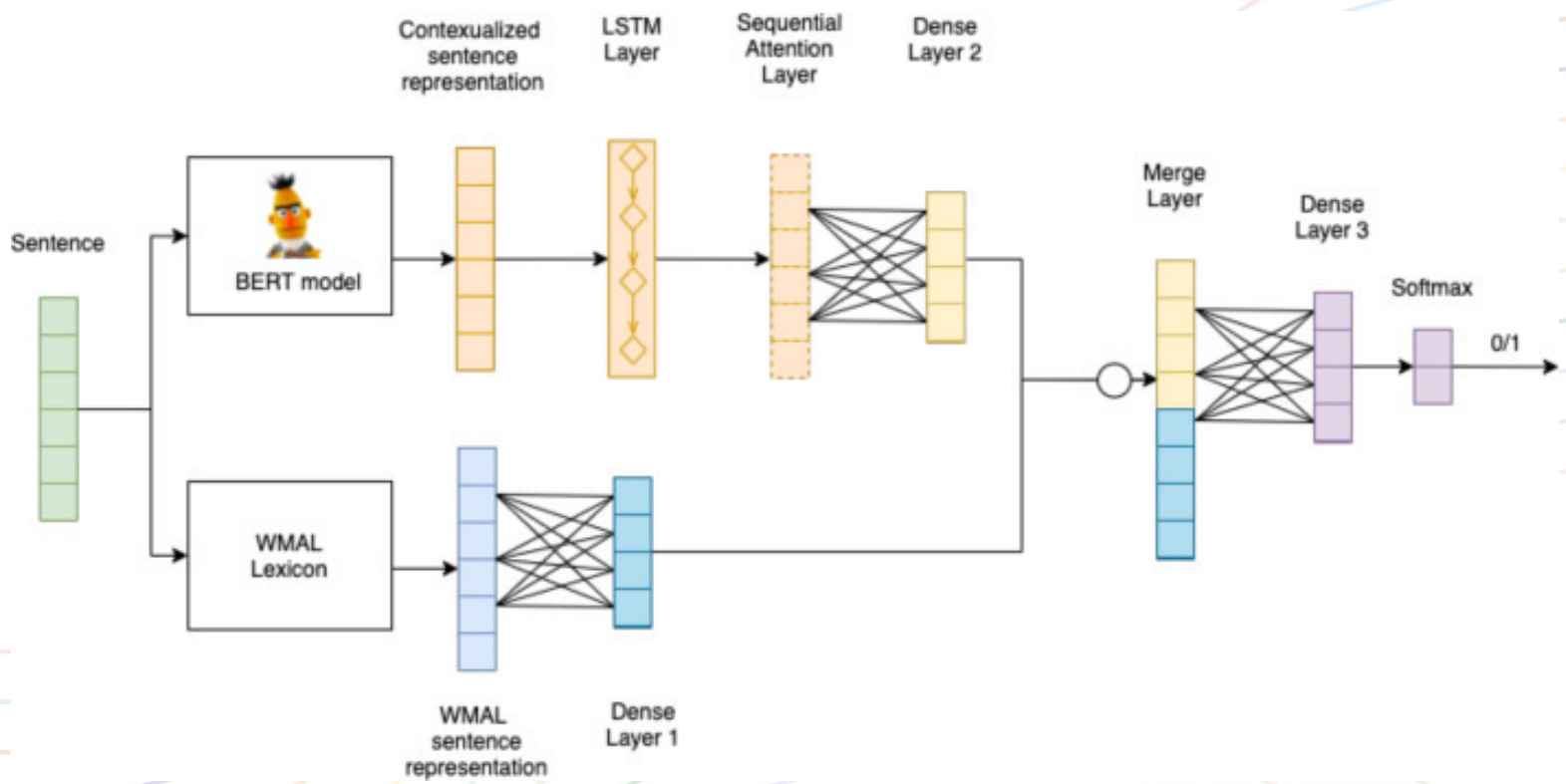
Other Datasets …

- *SENTIPOLC:* **7410** tweets for training and **2000** for testing
- *ABSITA*: **2365** for training and **1171** for testing

	# Train set	# Test set	Labels Pos	Labels Neg	Polarity
SENTIPOLC	7410	2000	0, 1	0, 1	Positive, negative, neutral, mixed
AGRITREND	/	1000	0, 1	0, 1	Positive, negative, neutral, mixed
ABSITA	2365	1171	0, 1	0, 1	Positive, negative, neutral, mixed

Polignano, M., Basile, V., Basile, P., Gabrieli, G., Vassallo, M., & Bosco, C. (2022). A hybrid lexicon-based and neural approach for explainable polarity detection. *Information Processing & Management*, *59*(5), 103058.

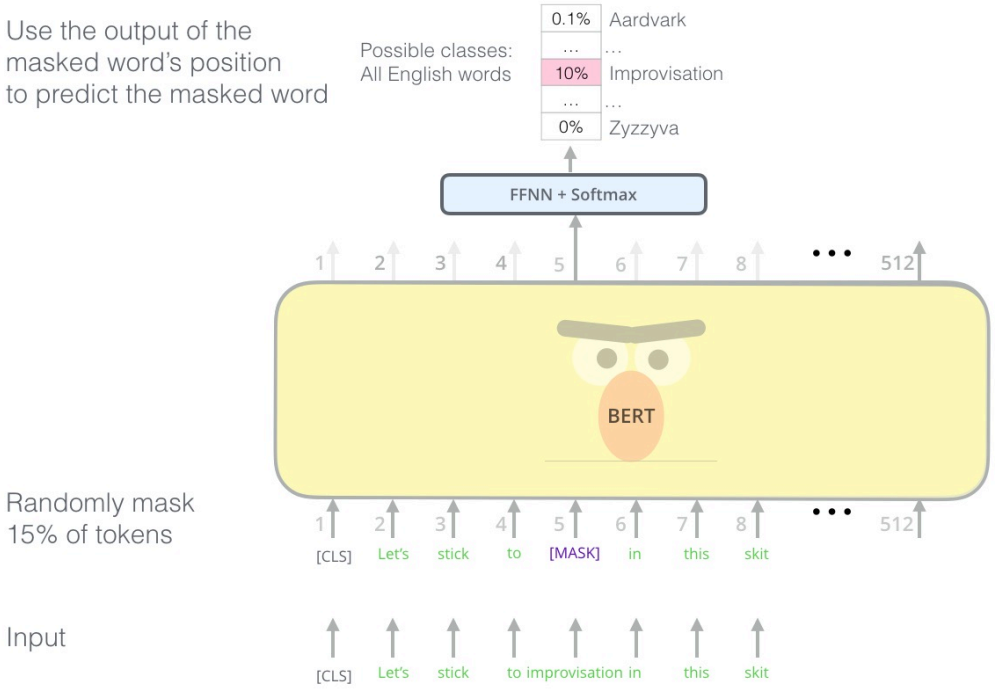
Deep Learning Architecture



Polignano, M., Basile, P., De Gemmis, M., Semeraro, G., & Basile, V. (2019). Alberto: Italian BERT language understanding model for NLP challenging tasks based on tweets. In *CEUR Workshop Proceedings* (Vol. 2481, pp. 1-6). CEUR.

AlBERTo: The Italian Language Understanding Model

AlBERTo wants to be the first Italian language understanding model to represent a style of writing of social networks, **Twitter** in particular, written in **Italian**.



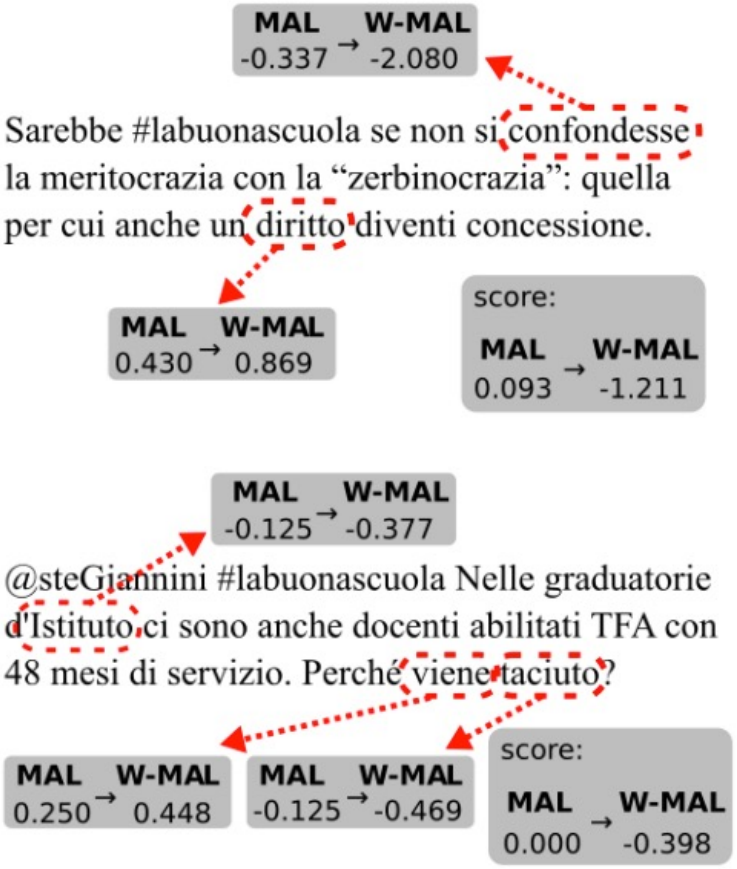
The **core** deep learning structure of BERT and AlBERTo is a **12x Transformer** Encoder, where for each input, a percentage of terms is **Masked** and then predicted for optimizing network weights in back-propagation.

https://github.com/marcopoli/AlBERTo-it

Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Vassallo M., Gabrieli G., Basile V., Bosco C. Polarity imbalance in lexicon-based sentiment analysis. Proceedings of the seventh Italian conference on computational linguistics, CLiC-It 2020, Bologna, Italy, March 1-3, 2021, CEUR workshop proceedings, vol. 2769, CEUR-WS.org (2020)

WMAL



Polignano, M., Basile, V., Basile, P., Gabrieli, G., Vassallo, M., & Bosco, C. (2022). A hybrid lexicon-based and neural approach for explainable polarity detection. *Information Processing & Management*, *59*(5), 103058.

Deep Learning Architecture

To sum up the key points about the process and the proposed model are:

- **Pre-processing**: removal of punctuation, non-alphanumeric elements, repeated space. Text transformed in lower case.
- **Generation of BERT embeddings**: Italian language BERT model, i.e.AlBERTo small version (L = 12, H = 748, A = 12, uncased).[3] Sum of last four layers. 128 tokens length. 128×768 shape.
- **Generation of WMAL embeddings**: 128 tokens length vectors with 0 if WMAL value not found for the corresponding token, the value otherwise.
- **Model**: Keras Functional API for a two-input model. Hyper-parameters of the different configurations evaluated are reported in the following.

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Deep Learning Architecture

In particular, we decided to use an *"Additive Layer"* between the output of the *"Sequential Attention Layer"* and the *"LSTM Layer"*. Then the *"Dense Layer 2"* has been kept for reducing the output the *"Sequential Attention Layer"* to 64 hidden-units.

Model	*AGRITREND*			*SENTIPOLC*			*ABSITA*		
	NEG	POS	AVG	NEG	POS	AVG	NEG	POS	AVG
Config 1	0.75761	0.36917	0.56339	0.72550	**0.86100**	0.79325	0.91631	0.91375	0.91503
Config 2	0.78093	0.42698	0.60395	0.74700	0.80750	0.77725	0.92400	0.90094	0.91247
Config 3	0.75963	0.39148	0.57555	0.75400	0.85300	0.80350	0.91887	0.91631	0.91759
Config 4	**0.79513**	**0.63083**	**0.71298**	**0.76400**	0.83100	**0.79750**	**0.92741**	**0.91887**	**0.92314**
Config 5	0.73935	0.53854	0.63894	0.71800	0.81950	0.76875	0.90350	0.89838	0.90094
AlBERTo	0.77521	0.61756	0.69638	0.74750	0.82440	0.78595	0.91973	0.91289	0.91631
no WMAL	0.78499	0.62880	0.70689	0.75820	0.83200	0.79510	0.90521	0.90863	0.90692

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Lexicon-driven classification explanation

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Post-Hoc Evaluation

The two concepts we have taken as reference over the WMAL model features:

- ***"contextual importance" (CI)*: How important for the prediction given by the model is the WMAL value assigned to the specific sentence token?".**
- ***"contextual utility" (CU)*** **:** "Compared to the possible variety of probability values obtainable for the WMAL value, how good is the one assigned? Is there a better WMAL value for obtaining the predicted class with a higher confidence score?".

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Post-Hoc Evaluation

We vary, for each sentence s_i of the test set, the WMAL value of each **j-esim term** of the sentence **100 times** and we look at the probability predictions of the model.

After this process we obtain ***s_len×100*** new sentences.

The new WMAL values were obtained from a Standard Gaussian distribution with mean given by the average of all the WMAL values in the training set and standard deviation 3xSigma, so that we had a 99.73% probability of obtaining new values in the WMAL original values range.

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Post-Hoc Evaluation

$$CI\left(s_{i,j}\right) = P_{max}\left(s_{i,j,k}, C_i\right) - P_{min}\left(s_{i,j,k}, C_i\right)$$

For each prediction class, if the variance is high then the value of WMAL strongly influences the prediction. So the term is **IMPORTANT**!

Contextual Importance

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Post-Hoc Evaluation

$$CU(s_i) = \frac{y_{i,j} - P_{min}(s_{i,j,k}, C_i)}{P_{max}(s_{i,j,k}, C_i) - P_{min}(s_{i,j,k}, C_i)}$$

For each prediction class, if the original Sentence Class Probability is the maximum already available then the term is **USEFUL**!

Contextual Utility

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Post-Hoc Evaluation

Explanatory Attention

olivicoltura <user> <user> pianificazione e visione strategica per il settore agricolae <url>

WMAL Attention

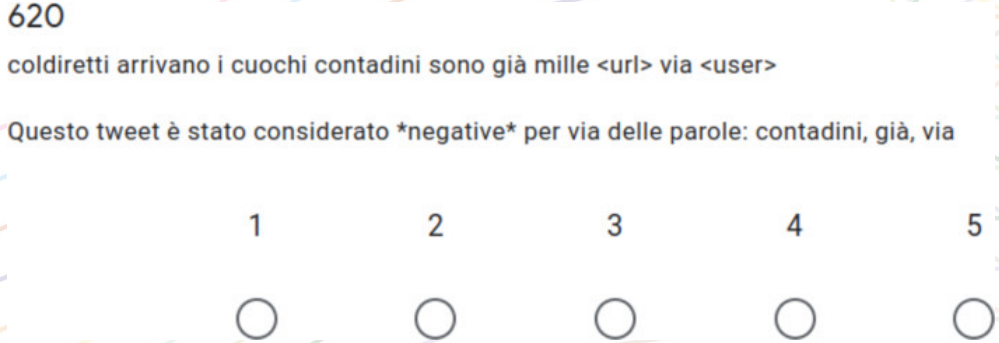
olivicoltura <user> <user> pianificazione e visione strategica per il settore agricolae <url>

- ***Attention Model Score*** x ***WMAL Values***

crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

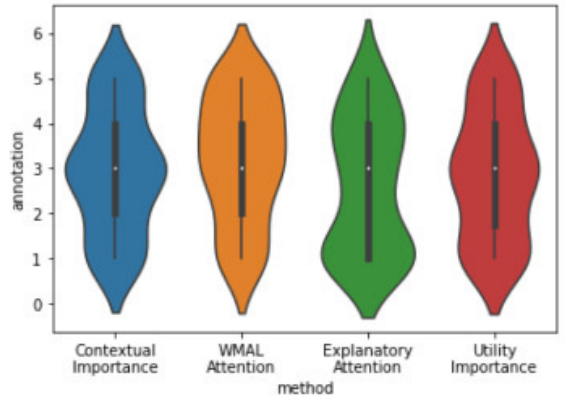
HUMAN Evaluation

- We selected a subset of **AGRITREND** made of **64 sentences**, 32 with positive polarity gold labels and 32 with negative polarity gold labels.
- For each of the 64 sentences, the text, a sentiment polarity label, three explanatory words (without numerical scores) and a rating scale are shown. The label shown is the prediction given by one of the four methods.

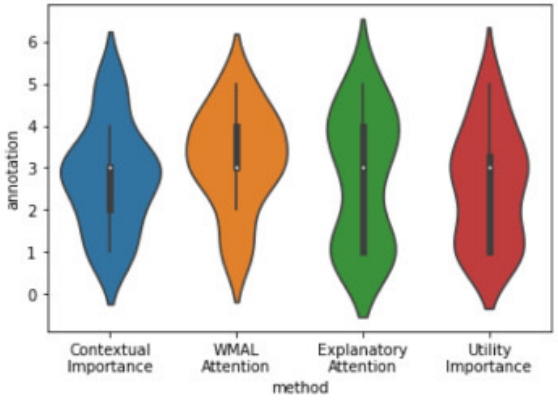


Results of the human evaluation in terms of average score given by the human judges (on a scale 1–5) and their standard deviation.

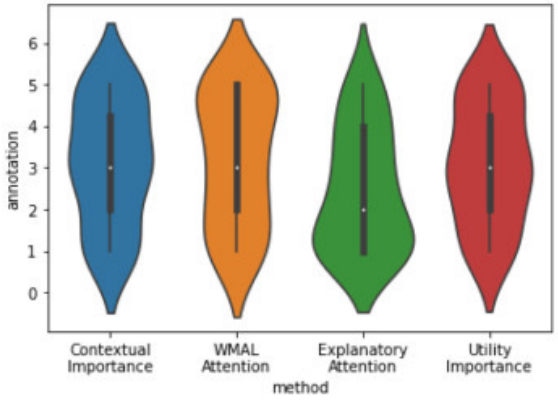
HUMAN Evaluation



(a) All instances



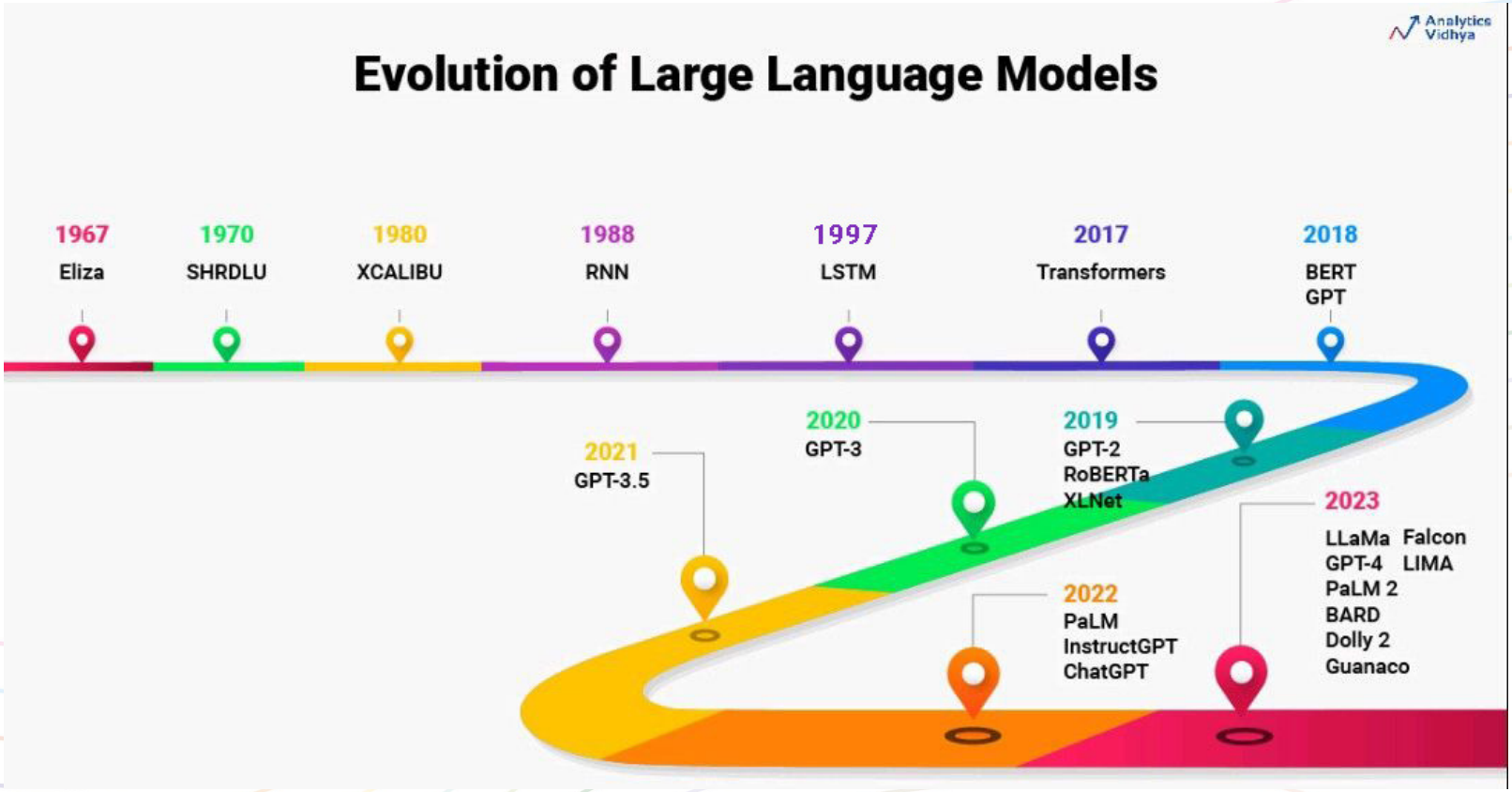
(b) Positive instances



(c) Negative instances

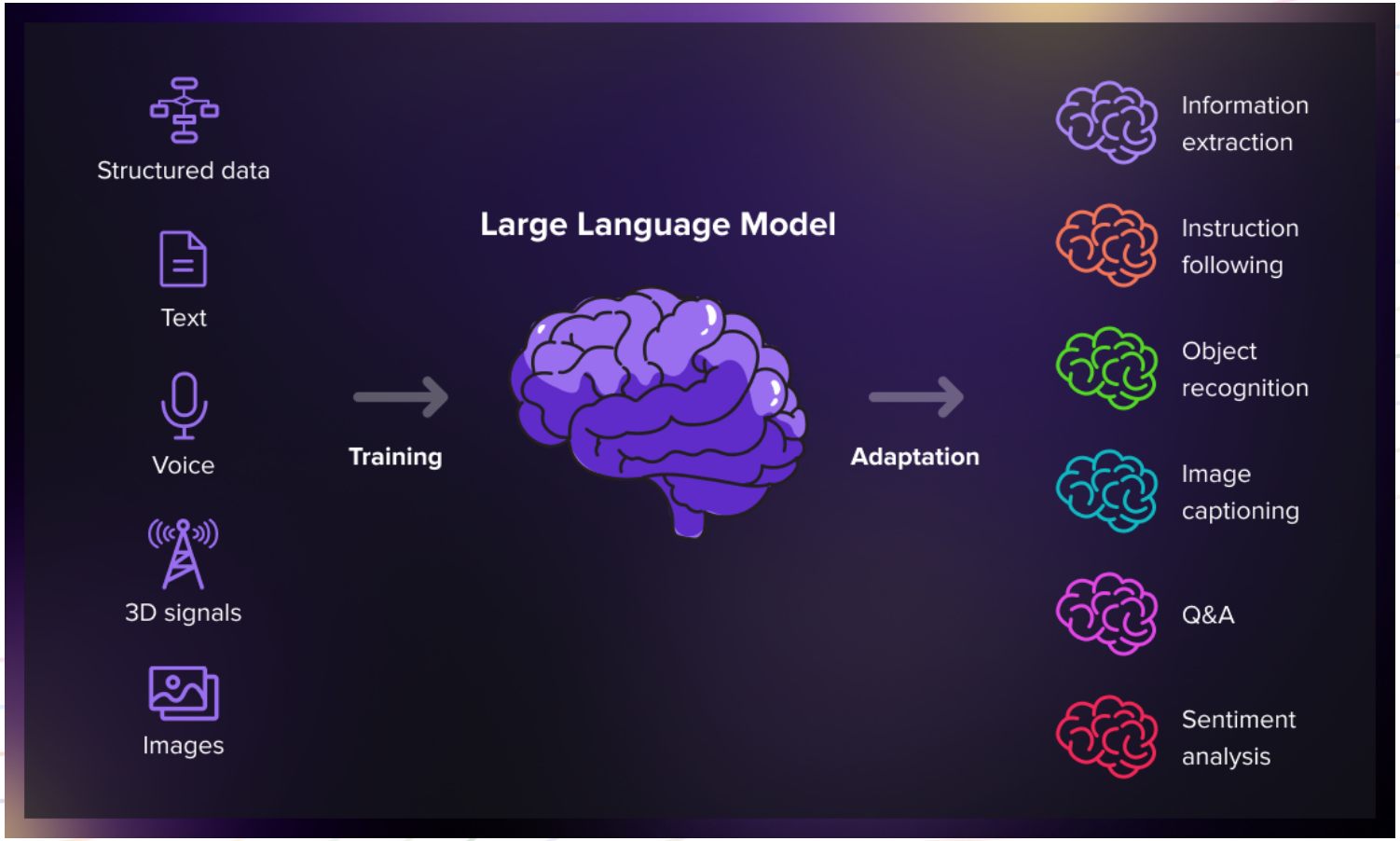
crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Future Directions …



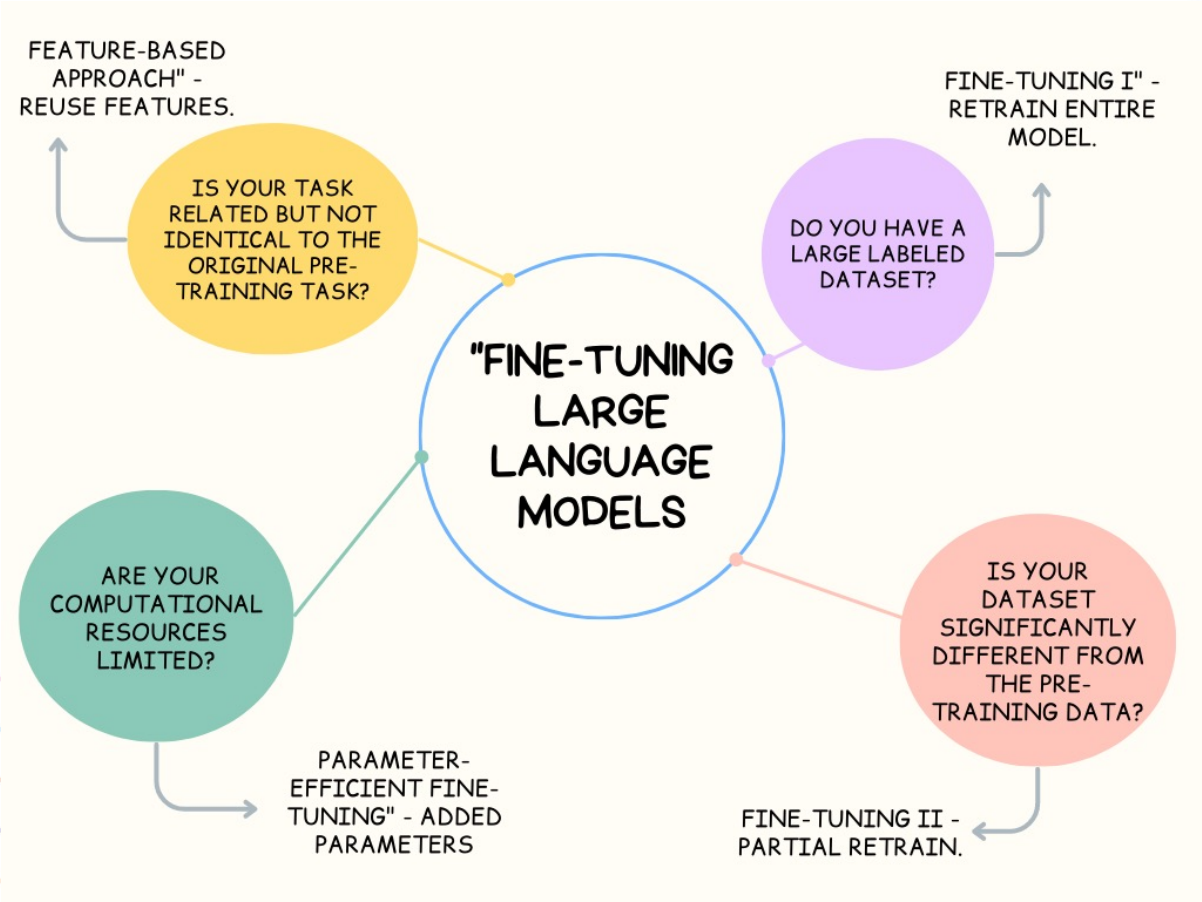
crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Future Directions …



crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Future Directions …



crea
Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Hromei, C. D., Croce, D., Basile, V., & Basili, R. (2023, September). ExtremITA at EVALITA 2023: Multi-task sustainable scaling to large language models at its extreme. In *Proceedings of the Eighth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2023), CEUR. org, Parma, Italy*.

Future Directions …

ExtremITA at EVALITA 2023: Multi-Task Sustainable Scaling to Large Language Models at its Extreme

Task	SubTask	Eval metric	extremIT5		extremITLLaMA		Best Competitor	
			Score	R	Score	R	Score	R
Emit	A	F1	0.5086	2	**0.6028**	**1**	0.4994	3
	B	F1	0.6331	2	**0.6459**	**1**	0.6184	3
EmotivITA	B	Pears Val	0.7080	4	**0.8110**	**1**	0.8110	2
		Pears Aro	0.4300		**0.6330**		0.6520	
		Pears Dom	0.5480		**0.6300**		0.6540	

crea
Consiglio per la ricerca in agricoltura
e l'analisi dell'economia agraria

Thank you!