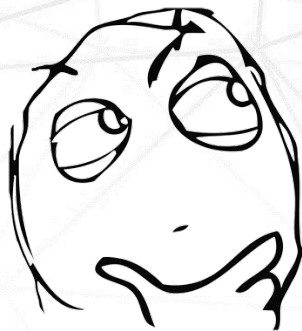


Introducció a Pytorch i NLPa

Assignatura: Models d'intel·ligència artificial

Professor: Ramon Mateo Navarro

Què és Pytorch?



PyTorch

Es tracta d'una biblioteca de codi obert dissenyada amb Python en ment i creada per a projectes d'aprenentatge automàtic. S'especialitza en diferenciació automàtica, càlculs de tensor i acceleració de GPU. Això ho fa especialment adequat per a aplicacions d'aprenentatge automàtic d'avantguarda com l'aprenentatge profund.

PyTorch és especialment popular entre els investigadors a causa de la personalització de Python. Crear capes de dades personalitzades i arquitectures de xarxa és especialment fàcil mitjançant Python.



PyTorch vs Tensorflow

	Keras	PyTorch	TensorFlow
API Level	High	Low	High and Low
Architecture	Simple, concise, readable	Complex, less readable	Not easy to use
Datasets	Smaller datasets	Large datasets, high performance	Large datasets, high performance

Debugging	Simple network, so debugging is not often needed	Good debugging capabilities	Difficult to conduct debugging
Does It Have Trained Models?	Yes	Yes	Yes
Popularity	Most popular	Third most popular	Second most popular
Speed	Slow, low performance	Fast, high-performance	Fast, high-performance

A practicar. Mireu el notebook Introducció PyTorch



Part 2: Què és el NLP?



Natural language processing

El Processament del Llenguatge Natural (NLP, per les seves sigles en anglès) és una branca de la intel·ligència artificial que se centra en la interacció entre els ordinadors i els humans a través del llenguatge natural. L'objectiu principal del NLP és permetre que els ordinadors processin i entenguin el llenguatge humà de manera que puguin realitzar tasques útils basades en aquesta comprensió.



Claus del NLP

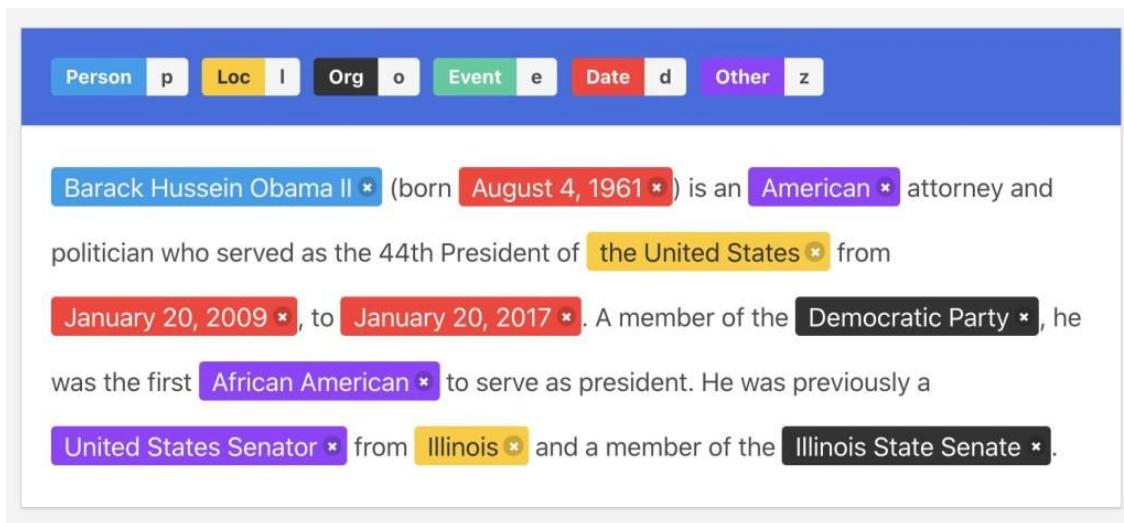
- **Comprensió i generació del llenguatge:** El NLP no només tracta de comprendre el text a nivell superficial sinó també d'interpretar les intencions, els sentiments i les implicacions subtils del llenguatge. A més, també inclou la capacitat de generar text coherent i contextualment adequat a partir de dades brutes.
- **Tècniques utilitzades:** El NLP utilitza diverses tècniques provinents de la lingüística, la ciència de dades i l'aprenentatge automàtic. Algunes d'aquestes tècniques inclouen l'anàlisi sintàctica, la modelització de temes, i l'aprenentatge profund (deep learning).

Tècniques: Tokenització i lemmatització

- La tokenització és el procés de dividir textos llargs en unitats més petites, com ara paraules o frases. Cada una d'aquestes unitats s'anomena "token".
- És un pas fonamental perquè moltes altres tècniques de NLP depenen de la manipulació de tokens individuals
- Lemmatització: Transforma la paraula al seu lema, que és la forma que trobaríem en un diccionari. Per exemple, "anava" es converteix en "anar".

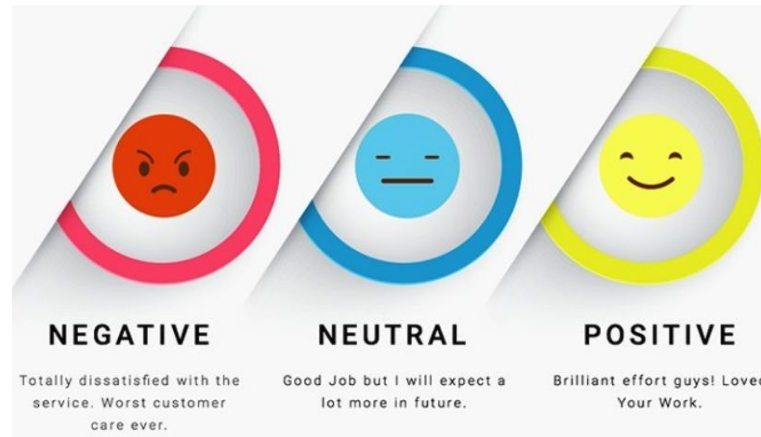
Tècniques: Reconeixença d'Entitats Nomenades (NER)

- La NER és una tècnica que identifica i classifica elements en categories predefinides (com noms de persones, organitzacions, ubicacions, etc.) dins d'un text.



Sentiment analysis

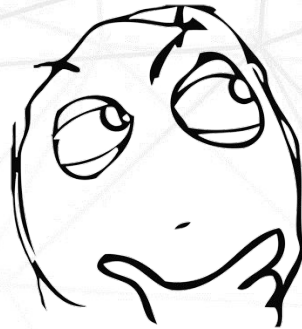
- L'anàlisi de sentiments determina l'actitud emocional subjacent d'un text, classificant-la com positiva, negativa o neutral



Resum

- Camps molt amplis: El NLP i el CV són branques profundes i establertes de la Intel·ligència Artificial amb una àmplia varietat de tècniques i aplicacions.
- Complexitat: Cada camp requereix un curs complet per abordar adequadament les seves múltiples tècniques i metodologies.
- Recomanació: Especialitzeu-vos en un àrea específica dins d'aquests camps expansius. Dominar un segment específic pot ser més manejable i productiu que intentar cobrir tot el camp.

Preguntes?



>
vedruna**pro**

Vall Terrassa

A programar



>
vedrunapro

Vall Terrassa

A programar



>
vedrunapro

Vall Terrassa