

HOT NLP HW3

VISUAL WSD

Cadoli Andrea 1827038

Clip

Clip Base

- Texts embeddings
- Image embeddings
- Similarity (mean of similarities)

	Accuracy
English	0.5053
Italian	0.1836
Farsi	0.07

Translation

Italian – English

- Transformer: MarianMT
- Model: 'Helsinki-NLP/opus-mt-it-en'

	Accuracy	Improvement
Italian	0.4295	+0.2459
Farsi	0.305	+0.235

Farsi – English

- Transformer: MT5
- Model: 'persiannlp/mt5-base-parsinlu-opus-translation-fa-en'

Not so accurate !

اسفنج دریایی -> Sorry
اسفنج -> Sea Sponge

Context

Gpt Neo, Prompt: 'Write a descriptive sentence about X'

- Neptune Statue -> The Neptune Statue is a sculpture of Neptune, the Roman god of the sea, located in the port of Naples, Italy
- Glasses -> The glasses are a type of glass

	Accuracy	Improvement
English	0.5702	+0.0649
Italian	0.446	+0.263
Farsi	0.33	+0.260

WSD

Simple

- First synset where Context appears in its definition

	Accuracy	Improv.
English	0.6306	+0.1253
Italian	0.495	+0.3144
Farsi	0.325	+0.255

Lesk

	Accuracy	Improv.
English	0.6328	+0.1275
Italian	0.495	+0.3144
Farsi	0.32	+0.250

Sentence Embeddings

- Context Embeddings - synset's definition embeddings similarity
- Embeddings with Sentence Transformer 'All-Mini-LM-L6'

	Accuracy	Improv.
English	0.6328	+0.1275
Italian	0.5016	+0.3155
Farsi	0.315	+0.245

Image Captioning

- Caption with Vit-gpt2-image-captioning
- Similarity caption-context with sentence embeddings

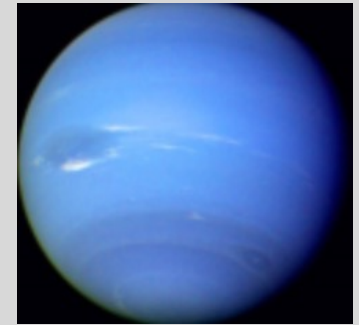
	Accuracy
English	0.2678
Italian	0.0393
Farsi	0.025

Vague



A statue of a man holding a sword
in front of a statue of a bird

Not accurate



A blue and white photo of a
blue and white frisbee

Results

	Clip base	Translation	GPT	WSD	Image Captioning
English	0.5053	-	0.5702	0.6328	0.2678
Italian	0.1836	0.4295	0.446	0.5016	0.0393
Farsi	0.07	0.305	0.33	0.315	0.025



THANK YOU 😊