

Intelligence Artificielle

Projet AI SPOTTER

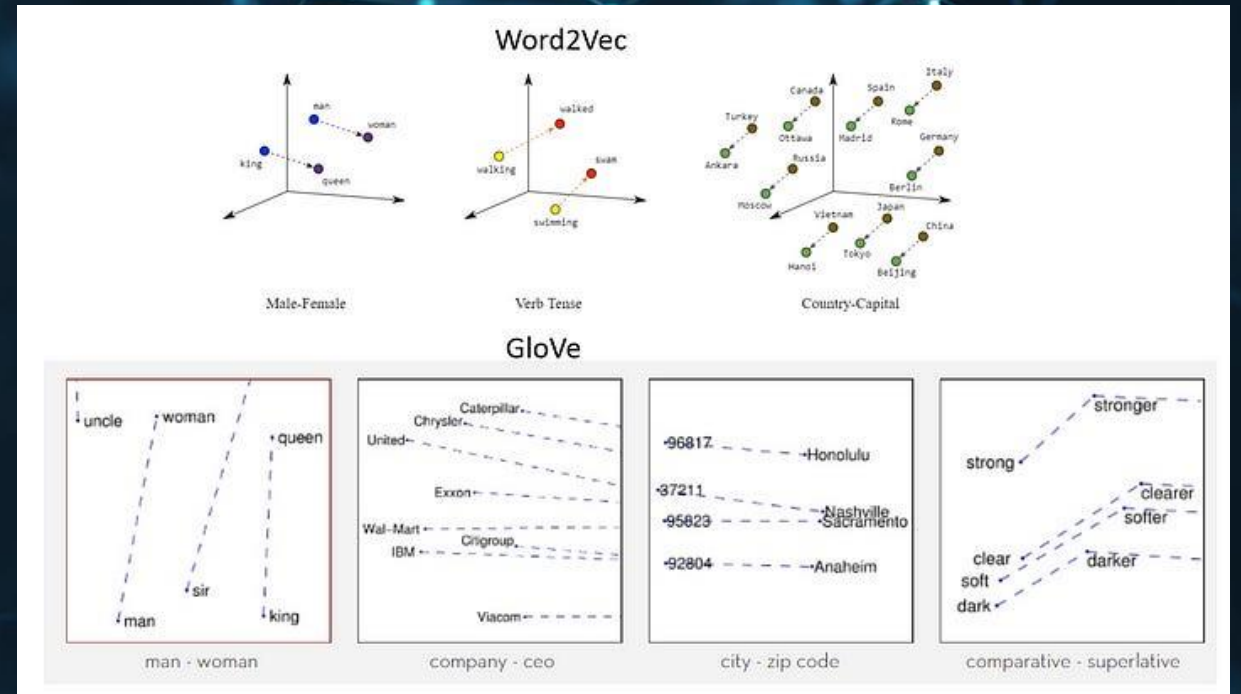
Étude des outils permettant l'analyse et la création d'un détecteur de textes générés par IA et utilisation des RCR et du ML pour répondre à une partie de la problématique.

L'analyse distributionnelle comme outil de développement du plongement lexical

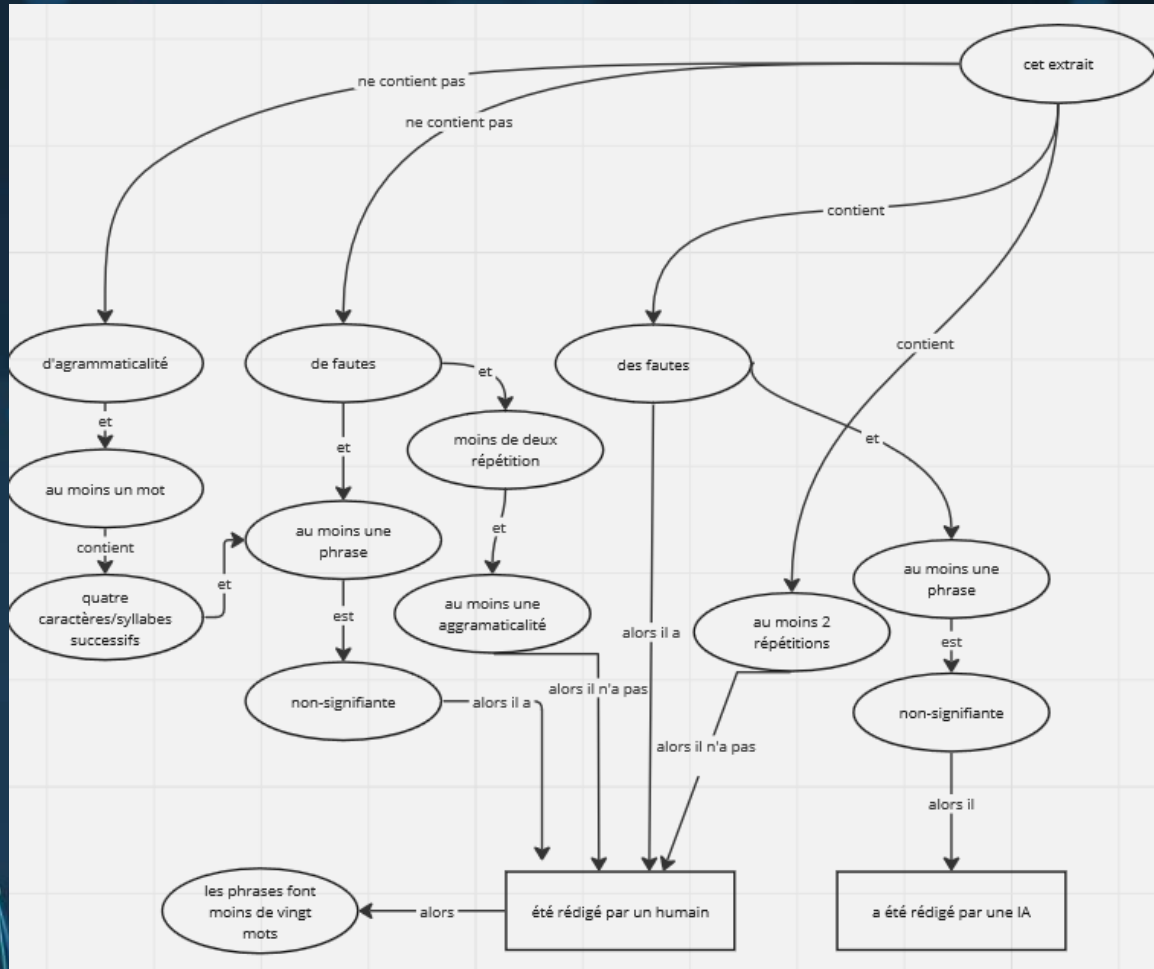
L'image à droite présente la représentation graphique de Word2Vec, il exprime la proximité entre deux mots par la proximité et la direction qui relie deux mots différents.

Sur les différents exemples, on voit que l'association se fait par catégorie et que dépendamment d'elle, les associations se font différemment.

Le fonctionnement derrière Word2Vec est l'utilisation à double couche, avec deux modèles possibles qui sont le CBOW : « Modèle de sacs de mots continus » et le modèle Skip-Gram qui se base sur un algorithme de matrices.



Représentation des connaissances



Graphe de connaissances du langage présenté dans le rapport.

Comme on peut le voir, on peut représenter notre langage de façon visuel afin de permettre de mieux analyser et comprendre notre problématique.

De manière similaire, il existe d'autres écritures et représentations différentes qui permettent d'être plus ou moins précis, d'être plus ou moins rigoureux (notamment sur le plan mathématique et logique par exemple)

Structure du projet

```
code/  
├── dataset/  
│   ├── dataset.csv  
│   └── dataset_modified.csv  
├── models/  
│   └── text_generation_model.h5  
├── scripts/  
│   └── script.py  
├── static/  
│   ├── css/  
│   │   └── style.css  
│   └── js/  
│       └── script.js  
├── templates/  
│   └── index.html  
├── README.md  
└── app.py
```

- ❑ Le fichier dataset.csv représente le jeu de données original, tandis que dataset_modified.csv est une version modifiée de ce jeu de données, où les valeurs 0 ont été remplacées par "Human" et les valeurs 1 par "AI".
- ❑ Le dossier models contient les modèles générés par script.py.
- ❑ Le fichier script.py est un script Python utilisé pour générer des modèles à partir d'un jeu de données en utilisant divers paramètres. Python a été choisi en raison de ses bibliothèques telles que TensorFlow et scikit-learn, qui facilitent la génération de modèles.
- ❑ Les dossiers static et templates contiennent les éléments front du projet.
- ❑ Enfin, app.py est le fichier principal du projet, l'API Flask.

Interface du projet

- **Interface utilisateur et API Flask**

- ❖ L'interface utilisateur a été conçue en utilisant les langages HTML, CSS et JavaScript pour une utilisation facile.
- ❖ L'intégration d'une API Flask en Python ajoute une couche fonctionnelle à l'interface. Elle permet de tester les textes en les envoyant au modèle, facilitant ainsi le processus d'analyse.

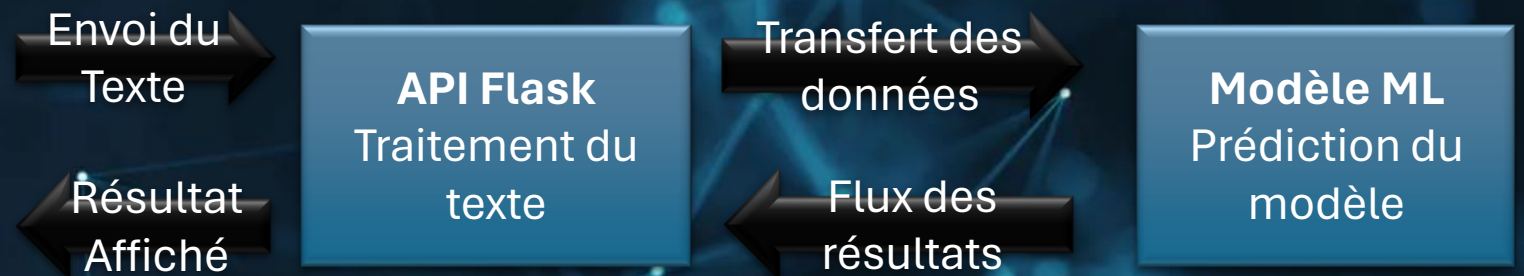


The screenshot shows a web interface titled "AI SPOTTER" in blue. Below the title, a subtitle reads "AI SPOTTER est un projet qui utilise un modèle d'intelligence artificielle pour effectuer des prédictions." There is a text input field with the placeholder "Saisissez votre texte ici" and a blue button labeled "Faire la prédiction". Above the input field, the text "Entrez votre texte :" is displayed.

Texte à Analyser

- **Acheminement des données**

- ❖ Les données de l'interface utilisateur sont acheminées vers l'API, où le texte est traité et le modèle appelé.
- ❖ La sortie générée par le modèle est ensuite renvoyée à l'API, qui l'affiche à l'utilisateur.



Configuration du modèle

- Le modèle généré par le script résulte de l'apprentissage automatique supervisé (machine learning) pour la classification de texte.

- Base de données:

Le base de données utilisé comprend 1000 textes en anglais, chaque entrée étant structurée avec un identifiant, un texte associé, et un entier indiquant si le texte a été généré par un être humain (0) ou par une intelligence artificielle (1). Voici un exemple d'une donnée:

```
1,"There are a variety of emerging applications for NLP, including the following; voice-controlled computer interfaces (such as in aircraft cockpits), programs that can assist with planning or other tasks,, more-realistic interactions with computer-controlled game characters, robots that interact with humans in various settings such as hospitals, automatic analysis or summarization of news stories and other text, intelligence and surveillance applications (analysis of communication, etc. ), data mining, creating consumer profiles, and other ecommerce applications, search-engine improvements, such as in determining relevancy",0
```

- Paramètres:

Nous avons configuré notre modèle en utilisant les paramètres suivants :

- ☐ Taille du Vocabulaire (vocab_size): 1000 - Limitant le modèle aux 1000 mots les plus fréquents.
- ☐ Dimension de l'Espace d'Embedding (embedding_dim): 50
- ☐ Longueur Maximale d'une Séquence (max_sequence_length): 1500

```
vocab_size = 1000
embedding_dim = 50
max_sequence_length = 1500
```


Entraînement du modèle

- **Séparation de l'ensemble de données**

Nous avons utilisé la méthode `train_test_split` pour diviser nos données en ensembles d'entraînement, de validation et de test.

Répartition des données :

- ✓ Ensemble d'entraînement : 70%
- ✓ Ensemble de validation : 15%
- ✓ Ensemble de test : 15%



- **Résultats de l'entraînement**

- Précision sur l'ensemble d'entraînement : ~99.02%
- Perte sur l'ensemble d'entraînement : ~0.0460
- Précision sur l'ensemble de validation : ~90.20%
- Perte sur l'ensemble de validation : ~0.3512

Le modèle a vraiment bien appris avec notre collection de textes, avec une précision d'environ 99.02% sur ceux qu'il avait vus auparavant. Cela signifie qu'il a bien compris les détails des textes générés par des humains et des IA en partie, grâce aux paramètres que nous avons choisis.

Cependant, il faut garder à l'esprit que ces résultats sont basés sur notre groupe de textes.

Exemples (1/2)

Nous avons demandé à ChatGPT de générer un paragraphe sur un domaine de l'informatique. Voici la réponse :

“The field of computer science is an ever-evolving landscape, characterized by constant innovation and technological advancements. In this dynamic realm, professionals explore the intricacies of algorithms, data structures, and programming languages to develop efficient and robust software solutions. The pursuit of artificial intelligence and machine learning has gained prominence, with researchers delving into the realms of neural networks and deep learning algorithms. Cybersecurity stands as a critical pillar, safeguarding digital assets from malicious threats. Cloud computing has revolutionized data storage and processing, offering scalable and flexible solutions. As we navigate the digital era, the fusion of hardware and software continues to shape the future of computing, promising unprecedented possibilities.”

Entrez votre texte :

processing, offering scalable and flexible solutions. As we navigate the digital era, the fusion of hardware and software continues to shape the future of computing, promising unprecedented possibilities.

Faire la prédiction

Prediction: AI

Exemples (2/2)

Nous avons rédigé un texte sur le domaine de l'informatique en mettant l'accent sur les répétitions et les fautes linguistiques :

“The computer science stuff is always changin', ya know? Like, there's these thingies called algorithms and data stuff, and we use 'em to make computer thingamajigs work. It's all about programmin' languages and makin' softwares that do cool stuff. People are super into makin' fake smart things with artificial brains and learnin' machines. Gotta watch out for them hacker dudes, 'cause cybersecurity is like, super important. Cloud thingies store data in the sky, and it's like magic or somethin'. Computers and software team up to make the future awesome, with lotsa potential and whatchamacallits.”

Entrez votre texte :

cybersecurity is like, super important. Cloud thingies store data in the sky, and it's like magic or somethin'. Computers and software team up to make the future awesome, with lotsa potential and whatchamacallits.

Faire la prédiction

Prediction: Human