

Algorithme de recrutement sans biais

JIHANE BENNIS | RACHEL GUIBERT | THAN VAN CON MARC PROJET OSY 40





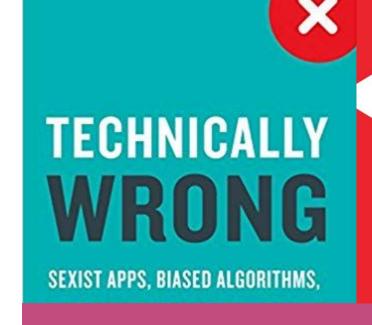
Le biais dans le Machine Learning



Applications au recrutement : des enjeux éthiques et économiques



Objectif long terme : Créer un algorithme "plus objectif" que des recruteurs humains



[A]rtificially Sex[I]st

Amazon's sexist hiring algorithm could still be better than a human

Maude Lavanchy







Client



Randstad entreprise spécialisée dans l'intérim

Volonté d'avoir un algorithme de matching job-CV Fourniture des CVs

1.Mise en correspondance offres-CV

2.Notre technique d'élimination du biais

3. Préparation des données textuelles pour l'entraînement

4. Entraînement de la boucle adversariale

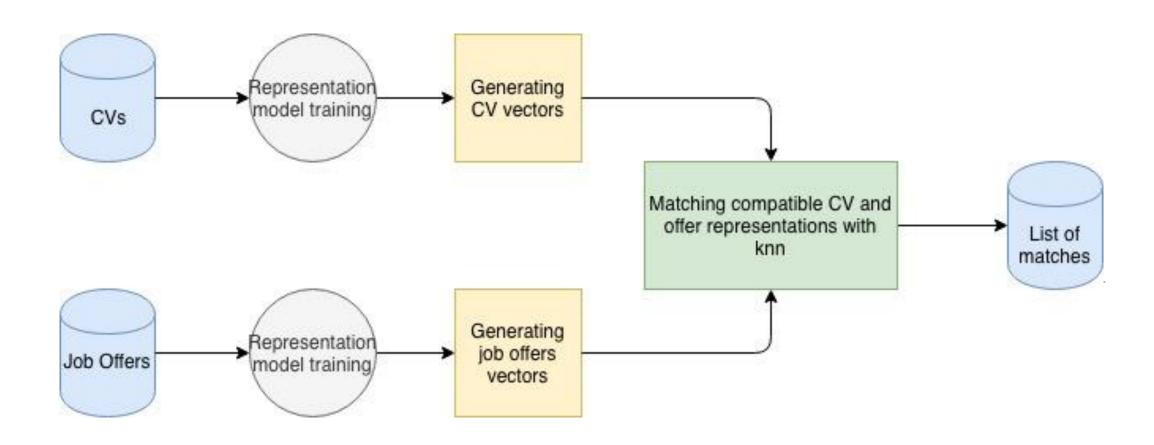
5. Résultats et limites de nos travaux





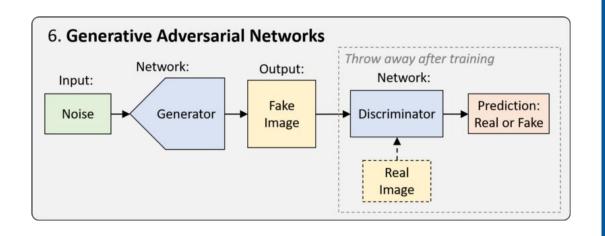


La mise en correspondance offres/CV





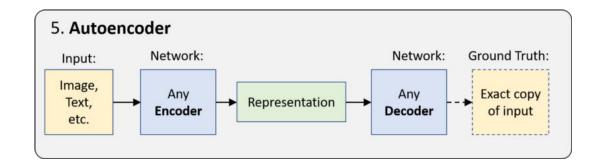


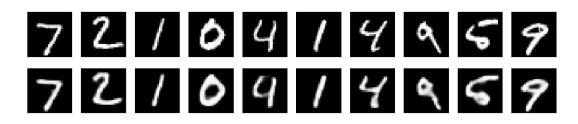




L'adversarial learning, notre première piste

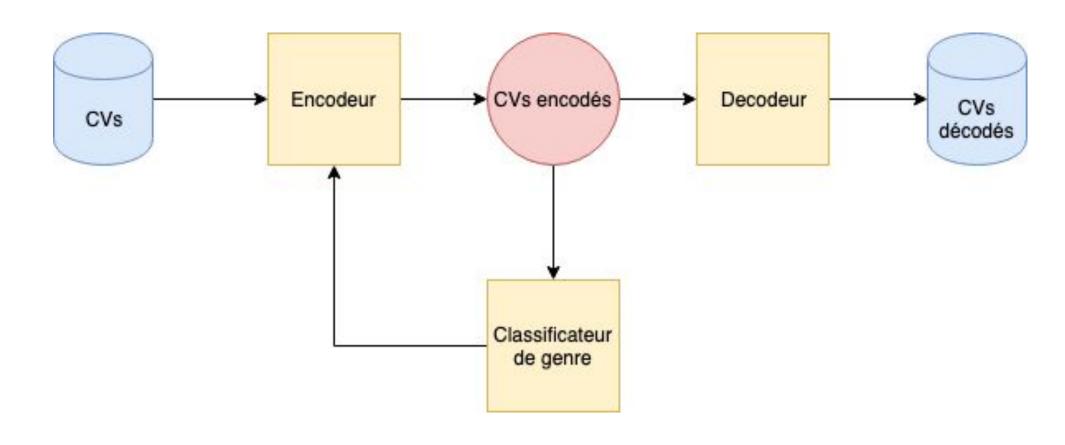






Auto-encoders pour garder l'information sur les compétences





Autoencoder adversarial pour répondre à notre objectif: une élimination du biais de genre sans perte d'information



Labellisation du genre



Robert Williamson

Product Engineer

PROFILE

Industry expert in Structural Engineering with excellent understanding of Drilling and Exploration techniques, Geology, and mechanical equipment used in underground and open pit mining operations. Safety focused with analytical approach to using numerical simulations to predict the behavior of underground excavations.

- Comprehensive knowledge of Finite Element Analysis techniques using a variety of software, as well as strong math and physics skills.
- Ensure smooth business operations by following projects from conceptualizing phase to completion, including system design, incorporation of customer needs, and technical support.

WORK EXPERIENCE



POWERTRAIN (FORMER INDUSTRIES INC.)

Developed structural finite element analysis(FEA) simulations and design optimizations. Managed file translations and drawings updates. Participated in product design on Honda projects. Developed solutions for FEA problems using variety of software applications. Organized structural FEA procedures for Ford flex plate group. Technical expert in stamping, mechanical, CAD and FEA issues. Developed new technologies in order to produce better, safer and less expensive products. Technical counselor on highly technical issues for the plants.

- Helped with CAD and FEA quoting for a stamped valve cover for Company resulting in project work of over \$1M.
- Saved over \$200K in tooling rework through numerical simulations.
- Created FEA simulations in the prototype phase and saved \$70K.

1995 - 2000 Product Engineer

PRODUCTS - CITY, STATE

Developed structural FEA simulations and design optimizations. Worked on-site in Detroit with Jeep Truck Engineering Group on Chrysler KJ 2002 program product design. (brought in jobs for the Company in value of over \$1M). Product Engineer tasked to follow up on a new product from its inception, going through subsequent CAD modifications, FEA simulations, to completion as a final product on the plant floor. Updated production plants with any changes for a specific product.

· Secured over \$1M in project work for the company.

EDUCATION

Bachelor of Science, Mechanical Engineering



Open Cast Mining and Quarrying - CESECO



4 +91 730 8100339

☑ robert.williams@visualcv.com

DETAILS

Date of Birth: 02.02.1961

Nationality: American

Gender: Male Marital Status: Married

Children: 2

CORE COMPETENCIES

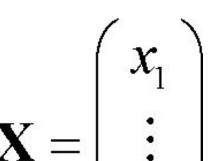
- Civil Engineering
- Structural Optimization Mining Stress Analysis
- Product Design
- Mechanical Engineering Drilling Geology
- Exploration Finite Element Analysis

TECHNICAL SUMMARY

- Microsoft Word
- Excel
- PowerPoint Publisher
- Project
- Outlook Access

Computer-Aided Design

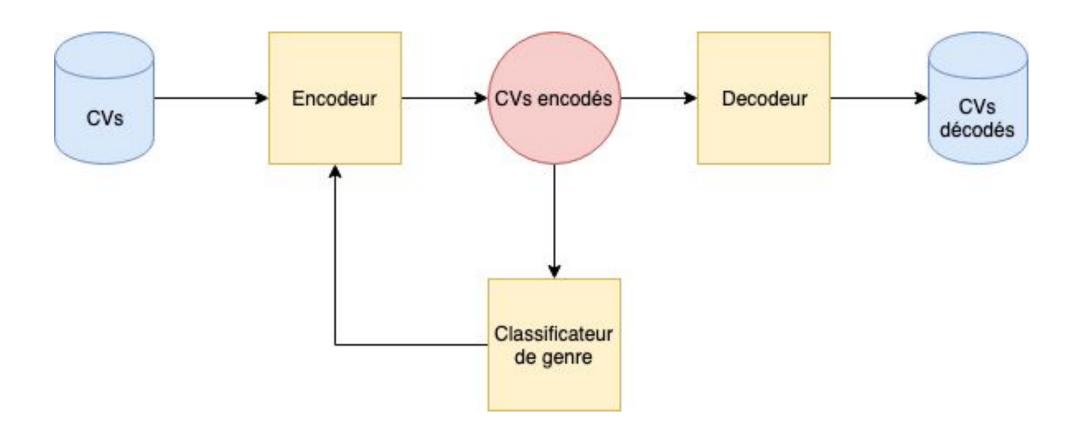
Tokenization







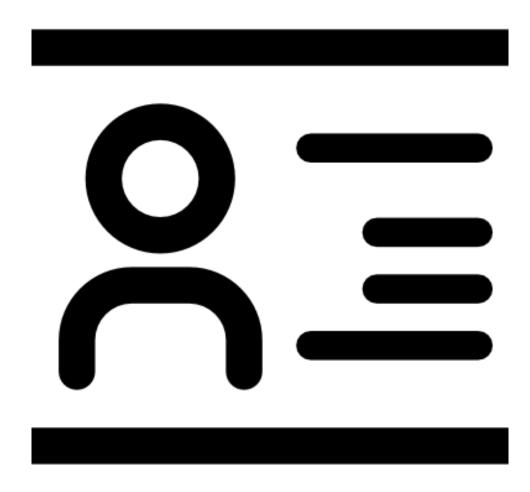




Notre boucle adversariale consiste en 3 MLPs avec une couche cachée pour chacun d'entre eux

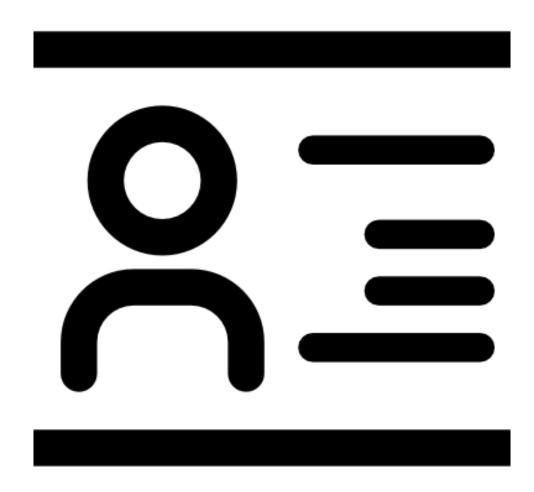
Carte d'identité du classificateur

- L'entrée: un vecteur de taille 1024
- L'activation : relu
- La sortie : un vecteur de taille 2 pour deux classes (femme et homme)
- La normalisation : softmax
- La fonction de coût : cross-entropie



Carte d'identité de l'encodeur/décodeur

- L'entrée : un vecteur de taille 1810
- L'encodeur : une couche cachée avec une sortie de taille 1024 (activation Relu)
- Le décodeur : une couche cachée avec une sortie de taille 1810 (activation Relu)
- La fonction de coût : Mean Squared Error





Technologies utilisées





Une fonction de coût globale pour notre réseau

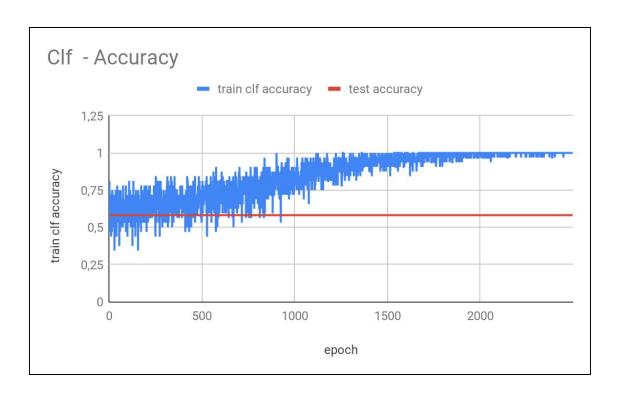
Adversarial_loss = Autoencoder_loss - \(\beta \) Classifier Loss

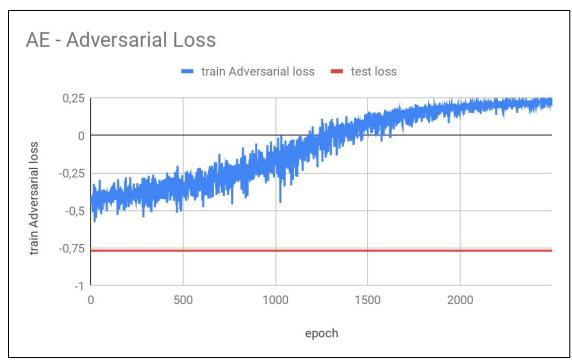
Objectif en minimisant Adversarial loss:

- Accuracy à 0.5→ Forcer Classifier Loss à augmenter [Oscillateur]
- **Conserver information AE** → Minimiser Auto-Encoder loss



Problématique du modèle





OverFitting

Adversarial non efficace



Paramètres | | | | | |

Taille du batch

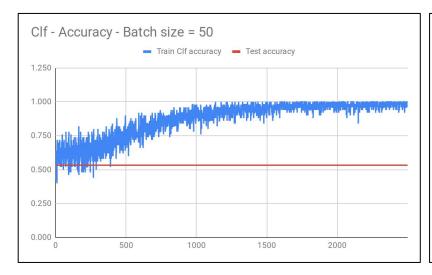
Bêta

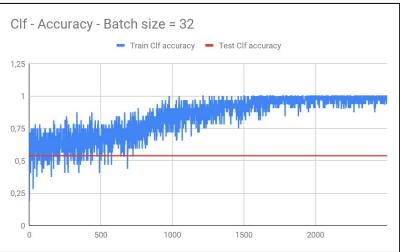
Pas d'apprentissage

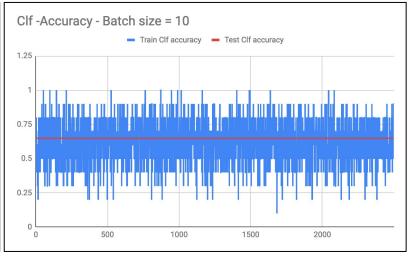
Alternance d'apprentissage

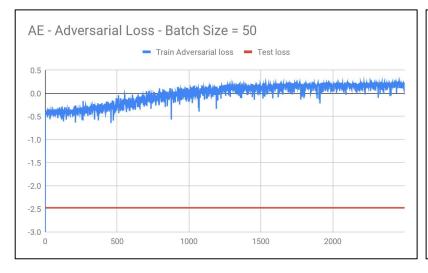
- Taille du batch

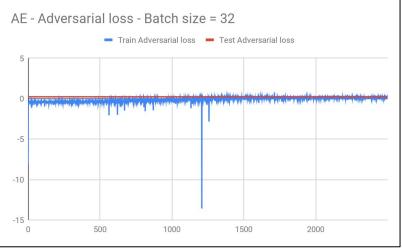


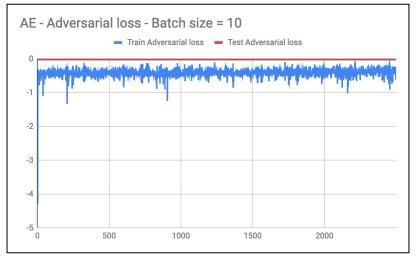




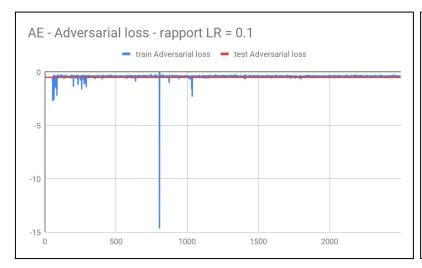


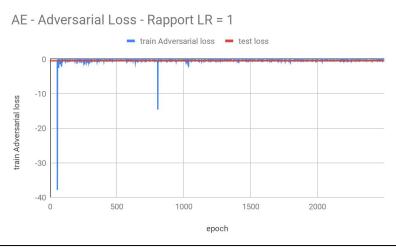


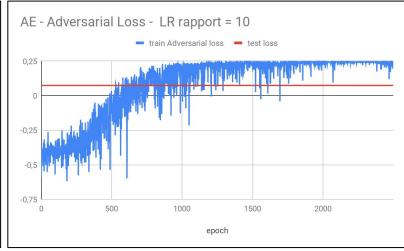


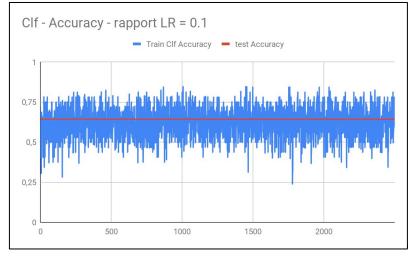


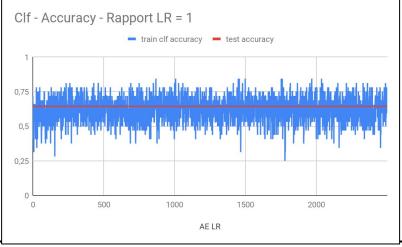
- Pas d'apprentissage: Rapport AE/Clf

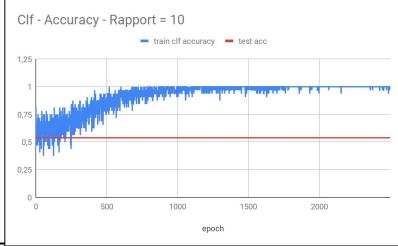






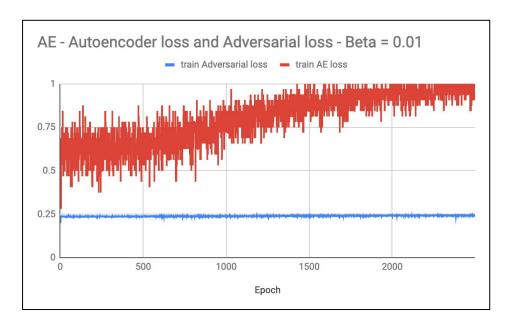


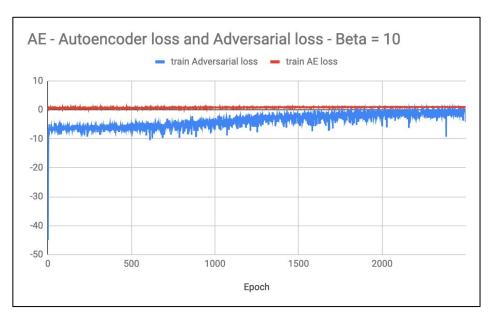


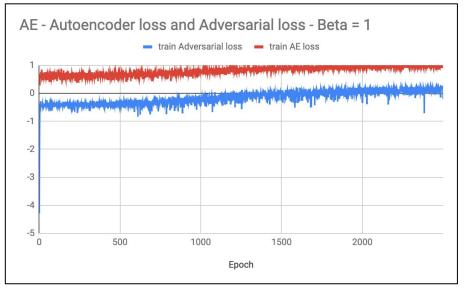


– Beta

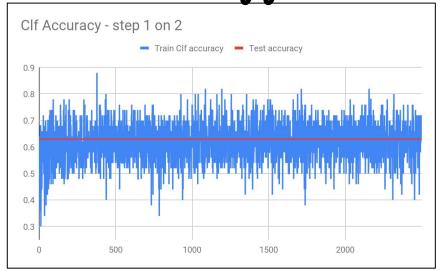


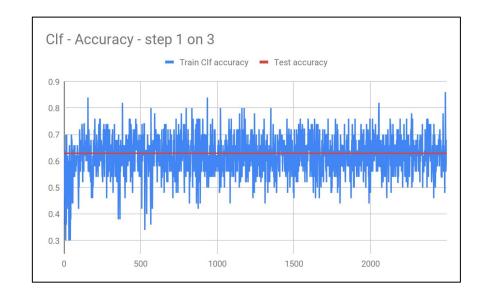


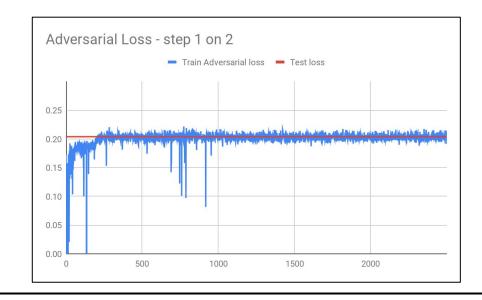


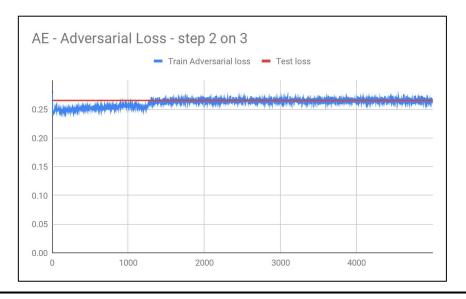


- Alternance d'apprentissage









Pistes d'amélioration

- Entrée
 - Différentes techniques de vectorisation (Onehot, TFid, Word2Vec)
- Architecture du modèle
 - Augmentation du nombre de couches
 - Réseaux de neurone différent (LSTM, CNN par exemples)
- Fonction de coûts différentes car objectif sur accuracy et non loss
- Hyperparamètre du modèle
 - Etude approfondie du paramètre Bêta
- Régularisation du modèle

Merci pour votre écoute

EST-CE QUE VOUS AVEZ DES QUESTIONS?

