

Objectif: Explorer, visualiser et classifier le jeu de données Fashion-MNIST.

- Le projet doit être réalisé en binôme. Merci de bien préciser les noms et prénoms de chaque membre du binôme et de mettre en copie votre binôme lors de l'envoi des scripts et de votre rapport.
 - Les résultats de vos analyses doivent être commentés et reprendre les différents notions vues en cours.
 - Vous devez me faire parvenir votre projet (scripts et rapport) pour le 19/12/2019 à 18h (*severine.affeldt@parisdescartes.fr*). Titre du message [MLSD-FI Deep Learning]
-

1. Réduction de dimension pour Fashion-MNIST avec une ACP

Faites un Analyse en Composantes Principales (ACP) du jeu de données Fashion-MNIST. Visualisez les résultats sur les différentes premières composantes. Commentez les clusters obtenus et la variance pour les composantes de la PCA.

Affichez un exemple d'image originale et reconstruite pour chaque digit.

2. Réduction de dimension pour Fashion-MNIST avec t-SNE

Utilisez l'algorithme t-SNE (t-distributed Stochastic Neighbor Embedding) ^{1 2} avec deux composantes pour faire une réduction de dimension de Fashion-MNIST. Etudiez les différents hyperparameters de l'algorithme, faites les varier et expliquez votre choix final ³.

Visualisez et commentez les clusters obtenus sur les deux composantes de t-SNE.

3. Réduction de dimension pour Fashion-MNIST avec deep autoencoder

Proposez une architecture et des hyperparamètres pour un autoencoder profond afin de traiter les images de Fashion-MNIST. Commentez vos choix d'architecture et de paramètres. Tracer la décroissance de la loss au cours de l'apprentissage.

Affichez un exemple d'image originale et reconstruite pour chaque digit.

4. Clustering dans un espace latent de petite dimension

Encodez les images de Fashion-MNIST à l'aide de votre autoencoders. Faites un k-means sur les images encodées, pour 10 clusters. Evaluer la performance de votre clustering, et commentez.

Appliquez l'algorithme t-SNE sur les images encodées. Visualisez et commentez les clusters obtenus sur les deux composantes de t-SNE.

¹<http://jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf>

²<https://scikit-learn.org/stable/modules/generated/sklearn.manifold.TSNE.html>

³<https://distill.pub/2016/misread-tsne/>

5. Ensemble clustering dans un espace latent de petite dimension

Encodez les images de Fashion-MNIST à l'aide de plusieurs autoencoders. Proposez une approche de clustering qui prend en compte l'ensemble des encodages. Evaluer la performance de votre *ensemble clustering*, et commentez.

Appliquez l'algorithme t-SNE sur les images encodées. Visualisez et commentez les clusters obtenus sur les deux composantes de t-SNE.