

Partiel IMDS4 08/12/2023 :

On considère le dataset qui donne le cours des actions d'AMD, Intel et Micron sur 3495 jours (fichiers AMD.npy, INTC.npy et Micron.npy à télécharger en utilisant np.load).

Préparer les données pour entraîner un modèle avec des couches LSTM. En entrée du modèle, on fournit le cours de bourse des 3 entreprises (3 features) sur une durée de 20 jours, et en sortie on souhaite avoir une prévision du cours de ces 3 actions les 4 jours suivants. Pour créer le dataset, la fenêtre sera décalée d'une journée entre chaque observation.

1. Quelle sera la taille du dataset sous la forme *[samples, window_size, features]* ? On justifiera sa réponse.
2. Préparer le dataset et le séparer en trois parties : train (60%), validation (20%) et test (20%).
3. Construire le modèle en utilisant des couches LSTM et la MSE comme fonction coût. Le modèle comprendra une entrée devant accueillir des données sous la forme *[samples, window_size, features]* et 3 sorties : une pour chaque action. Chaque sortie devra prédire le cours d'une action sur 4 jours consécutifs. Combien de neurones doit comporter chacune de ces sorties ?
4. Faire une recherche d'hyperparamètres en utilisant les données de validation.
5. Entraîner le modèle final.
6. Pour chacune des actions, représenter sur un graphique le cours exact et le cours prédit par le modèle pour les données de validation+test.