

ARTICLE : PLANT DISEASES DETECTION AND CLASSIFICATION BASED ON DEEP LEARNING

Proposé par : M. MBOUNA

Étudié par :

DJIEMBOU Victor¹

TANEKEU Bill¹

ROBNDOH Mardochée¹

MEGNA Ibrahim¹

¹Supervisé par : Dr. AMINOU Halidou

29 mai 2024



1 Questions devoir 1

2 Questions devoir 2

3 Questions devoir 3

Compréhension du sujet de recherche

Q1 : Pourquoi prédire le risque de crédit bancaire ?

Pour limiter le défaut de crédit car le crédit bancaire est l'une des meilleures sources de revenu pour les banques mais expose à des risques de non remboursement.

Q2 : Quel est le problème à résoudre dans le contexte du crédit scoring ?

Trouver un moyen efficace d'évaluer la probabilité associée au remboursement d'un emprunt.

Q3 : Quelles sont les contraintes classiques sur les données de credit scoring ?

- Données déséquilibrés et peu représentatives,
- Le défaut de crédit n'a pas la même valeur pour tous les crédits.

Sous-problème d'extraction de descripteurs

Q1 : Comment construire un graphe multicouches dans le contexte du credit scoring ?

Il s'agit d'un graphe multicouches où

- chaque emprunteur a autant de nœud qu'il y a de couches considérées
- les nœuds de chaque emprunteur sont tous reliés les uns aux autres
- chaque attribut d'une dimension a un nœud associé
- si un emprunteur est décrit par un attribut dans une dimension donnée, alors le nœud emprunteur de cette dimension est relié au nœud attribut associé
- la navigation d'une couche à une autre se fait en passant par les nœuds emprunteurs des différentes couches

Sous-problème d'extraction de descripteurs

Q2: Comment extraire les nouveaux descripteurs du graphe multicouches ?

Les nouveaux descripteurs sont extraits en exécutant des algorithmes d'analyse de lien dans les graphes comme le PageRank qui fourni après analyse l'important de chaque noeud dans le graphe.

Q3: Quels sont les descripteurs extraits du graphe multicouches ?

- Dégrée d'un noeud : le nombre de noeuds emprunt qui partagent les mêmes descriptions dans les différentes couches
- Score de PageRank d'un noeud : l'information de PageRank d'un noeud après exécution de l'algorithme de PageRank sur le graphe.

Sous-problème d'extraction de descripteurs

Q4: Comment est-ce que les descripteurs extraits des graphes multicouches sont utilisés pour résoudre le problème du credit scoring ?

- Ils sont ajoutés en entré des jeux de données classic pour créer de nouvelles dimensions descriptives
- en espérant qu'elle permettent d'améliorer les performance métriques des modèles de machine learning
- et qu'ils soient des sources pertinentes pour les sortis des modèles.

Sous-problème d'extraction de descripteurs

Q5 : Comment apprécier l'impact de l'ajout des nouveaux descripteurs ?

- Performances :
 - Entrainer sur le jeu de données sans descripteurs
 - puis sur le jeu de données ajouté des descripteurs
 - comparer les performances métriques des modèles formés.
- Pertinence :
 - Plotter les courbes SHAP pour chaque modèles
 - et évaluer la présences des descripteur extraits dans le top k des meilleurs caractéristiques qui explique la sortie des modèles.

Implementation du sous-problème d'extraction de descripteurs.

Q1 : Citer deux types de descripteurs extraits des graphes ?

- les statistiques de degré (degré moyen, distribution des degrés)
- les mesures de centralité (centralité de degré, de proximité, d'intermédiarité)
- les caractéristiques topologiques (coefficient de clustering, diamètre, modularité)

Implementation du sous-problème d'extraction de descripteurs.

Q2: Quel est principe du Linear discriminate analysis ?

- **Réduction de dimensionnalité** : projeter les données d'entrée dans un sous-espace de plus faible dimension tout en préservant au mieux la séparabilité des classes tel qu'il maximiser la séparation entre les classes tout en minimisant la variance intra-classe.
- **Recherche d'un hyperplan optimal de séparation** : défini par un vecteur normal qui maximise le rapport entre la variance inter-classe et la variance intra-classe
- **Classification** : classer de nouvelles observations en les projetant sur cet hyperplan et en les assignant à la classe la plus proche

Implementation du sous-problème d'extraction de descripteurs.

Q3 : expliquer les 3 mots clés du principe de fonctionnement de l'algorithme du PageRank

- **Récursivité** : calcule de manière récursive l'importance (ou le "rang") d'un noeud en se basant sur les liens entrants provenant d'autres noeuds
- **Influence** : le rang d'un noeud dépend de l'influence des noeuds qui y pointent. Plus un noeud est influent (a un rang élevé), plus il transmet de l'importance aux noeuds qu'il référence
- **Itératif** : le calcul se fait de manière itérative jusqu'à convergence, permettant d'obtenir une estimation stable de l'importance relative de chaque noeud dans l'ensemble du graphe

Implementation du sous-problème d'extraction de descripteurs.

Q4: Q'est-ce-que XAI?

L'intelligence artificielle explicable (ou XAI) est un ensemble de processus et de méthodes qui permettent aux utilisateurs humains de comprendre et de faire confiance aux résultats créés par les algorithmes de machine learning.

Implementation du sous-problème d'extraction de descripteurs.

Q5 : Quels sont selon le NIST (Le National Institute of Standards and Technology), les quatres principes qui régissent le XAI ?

- **Explication** : Les systèmes fournissent des preuves ou des motifs qui accompagnent tous les résultats
- **Compréhensible** : Les systèmes fournissent des explications que les utilisateurs peuvent comprendre
- **Précision des explications** : Les explications traduisent parfaitement le processus du système pour produire les résultats
- **Limites des connaissances** : Le système ne fonctionne que dans les conditions prévues ou lorsque ses résultats ont atteint un niveau de confiance suffisant

Merci de votre attention

Victor DJIEMBOU¹

¹ *Université de Yaoundé I,
Faculté des Sciences,
Département d'Informatique,
Étudiant Master SD*