Fiche projet en R & D Sciences des données - IA

Renseignements généraux

OFFRE:

- Étudiant au doctorat ou à la maîtrise
- 4 mois temps partiel ou temps plein (Environ 225 h)
- 10 000\$ en bourse

Été 2023

Nom et prénom du responsable en entreprise : Chiranjib Chaudhuri

Courriel du responsable en entreprise : chiranjib.chaudhuri@geosapiens.ca + en CC. entrepreneuriat@ivado.ca

Titre du stage : Modèle d'apprentissage profond pour prédire la distribution des débits fluviaux extrêmes au Canada et aux États-Unis.

Mots clés: Régression régionale, apprentissage profond probabiliste, analyse de la fréquence des inondations

Description de l'entreprise:

Geosapiens est une entreprise spécialisée dans la modélisation des risques d'inondation. Nous développons des solutions en ligne pour aider les citoyens, les gouvernements et les compagnies d'assurance à mieux comprendre les risques d'inondation. Notre mission est d'augmenter la résilience de notre société face aux inondations, et donc de réduire les impacts sociaux, économiques et environnementaux. Nous sommes une entreprise innovante et en forte croissance. Nous visons à déployer nos solutions à travers l'Amérique du Nord dans les deux prochaines années, et nous sommes à la recherche de talents motivés pour nous aider à atteindre notre mission et nos objectifs.

Description du projet:

Les débits de période de retour sont utilisés dans une variété de secteurs, y compris la gestion de l'environnement et la planification des inondations. La plupart des bassins n'étant pas jaugés, ces processus sont souvent réalisés sans utiliser les débits quantiles observés. Dans cette situation, les algorithmes de régionalisation sont fréquemment utilisés pour prévoir la distribution des débits à l'aide de données provenant de bassins jaugés. Le développement de modèles de régression qui relient les débits observés à des facteurs indépendants dérivés des données physiques et météorologiques du bassin est un type fréquent de méthode de régionalisation. Les modèles hydrologiques basés sur des paramètres estimés et des données de forcage climatique peuvent être utilisés pour calculer les débits percentiles. Cependant, la facilité avec laquelle les modèles de régression peuvent être utilisés pour prévoir les débits percentiles dans les bassins non jaugés constitue une opportunité. Pour les très grandes zones de recherche présentant une variance considérable des débits, les modèles de régression sont connus pour être peu performants. Les progrès récents des modèles d'apprentissage automatique nous donnent l'occasion de créer un modèle unifié pour le Canada et les États-Unis. Objectif: 1. identification d'un ensemble de facteurs indépendants, par exemple l'emplacement de la station, la zone en amont, le lulc, les paramètres du sol, les variables topographiques, les paramètres climatiques durables, etc. 2. Créer un réseau de densité mixte (MDN) qui peut relier les paramètres indépendants à la distribution du débit aux jauges. 3. Estimer le RP ou la CDF du débit à l'aide du MDN aux jauges et les comparer/modifier par rapport aux estimations du gouvernement. 4. Prédire (déduire) la distribution à tous les endroits pour le Canada et les États-Unis.

Les compétences recherchés:

- Apprentissage automatique
- Apprentissage en profondeur
- Apprentissage par renforcement
- Statistiques et analyse de données

- Visualisation des données
- Exploration de données
- Python ou autre langage de programmation

R & D Project Sheet Data Science - Al

Informations

OFFER:

- Ph.D. or Master's student.
- 4 months part time or full time (Approximately 225 h)
- 10 000\$ in scholarship

Summer 2023

Person in charge at the company: Chiranjib Chaudhuri

Emails: chiranjib.chaudhuri@geosapiens.ca + en CC. entrepreneuriat@ivado.ca

Title of the project: Deep learning model to predict distribution of extreme river discharge over Canada and US.

Key words:

Regional regression, probabilistic deep learning, flood frequency analysis

Company description:

<u>Geosapiens</u> is a company specialized in flood risk modeling. We develop web-based solutions to help citizens, governments and insurance companies to have a better understanding of flood risks. Our mission is to increase our society's resilience in the face of flooding, and therefore, to reduce the social, economic and environmental impacts. We are an innovative and a fast-growing company. We aim to deploy our solutions across North America in the next two years, and we are looking for talents that are motivated to help us achieve our mission and goals.

Project description:

Return period flows are employed in a variety of sectors, including environmental management and flood planning. Because most basins are ungauged, these processes are frequently carried out without the use of observed quantile flows. In this situation, regionalization algorithms are frequently used to forecast flow distribution using data from gauged basins. The development of regression models that relate observed flows to independent factors derived from physical and meteorological basin data is a frequent sort of regionalization method. Hydrologic models based on estimated parameters and climatic forcing data can be used to calculate percentile flows as an alternative. However, the ease with which regression models may be used to forecast percentile flows in ungauged basins presents an opportunity. For very large research areas with a considerable variance in flows regression models are known to perform poorly. The recent advancement of machine learning model presents us an opportunity to create a unified model for Canada and US. Objective: 1. Identification of a set of independent factors, for example location of the station, upstream area, lulc, soil parameters, topographic variables, sustained climate parameters etc., which are relevant in modifying the annual maximum river discharge. 2. Create a Mixed Density Network (MDN) which can relate the independent parameters to the distribution of the discharge at gauges. 3. Estimate the RP or CDF of the flow using the MDN at the gauges and compare/modify them against the govt. estimates. 4. Predict (infer) the distribution at all the locations for Canada and US.

Skills:

- Machine learning
- Deep learning

- Reinforcement learning Statistics and data analysis Data visualization
- Data mining
- Python or other programming language