

Artificial Cohort Generation for Clinical Research at CHUV using Machine Learning

(30.09.2019)

Stage N° 27494

Format Stage Master, Projet de Master (diplôme)

Nombre de places 1

Durée 4 - 6 mois

Périodes Période 1 : février - septembre (2020)

Informations sur l'entreprise

Entreprise (mère) CHUV Centre hospitalier universitaire vaudois

Entreprise (labo, dpt, ...)

Adresse Rue du Bugnon 21

URL

Localité 1011 Lausanne, Suisse

Informations sur le contact

Contact Despraz Jérémie

Téléphone professionnel 021 314 46 46

Titre

Téléphone mobile

Email jeremie.despraz@chuv.ch

Description du travail

Description et objectifs

The Lausanne University Hospital (CHUV) is a key player in medical care and biomedical research both at national and international levels. The mission of the Data Science team within CHUV's Department of Information Systems, is to foster the adoption and use of data science and innovative tools in medical informatics inside the hospital to significantly improve biomedical research and hospital key processes.

To achieve this ambitious goal, one of the core tasks of the group is to develop the necessary IT infrastructure for providing physicians and scientists with the means to smoothly access and use, for their clinical research projects, the massive volume of clinical data that is generated every day within the hospital. Our current and future challenges lie at the intersection of big data, medical informatics, data protection and artificial intelligence.

Collaboration and data sharing have become core elements of biomedical research. Protective measures like anonymization and synthetic dataset generation are paramount for gaining patients' trust and bringing together data from healthcare and research to conduct advanced studies.

To leverage the ability of the Data Science group to provide researchers with high quality data while complying with current privacy guidelines on data protection, you will implement a tool able to generate artificial data points that are statistically identical to the CHUV datasets. This tool will be constructed using state of the art Machine Learning methods.

In order to achieve this goal, you will:

- 1) explore existing methods and tools to generate synthetic clinical data
- 2) test the selected approaches on simple datasets
- 3) explore the extension capabilities of the tool to more complex datasets and data types
- 4) assess the compliance of the generated data with data privacy guidelines such as HIPAA
- 5) implement a robust version of the final method and deploy it for potential CHUV end-users

Compétences requises

- Knowledge of state-of-the-art Machine Learning methods and statistics
- Strong Linux background
- Good proficiency in Python and ML related modules (jupyter, scikit-learn, pandas, numpy, etc) any other lower level programming language is a plus
- Knowledge of free machine learning software libraries such as Tensorflow, Keras, or pyTorch
- Interest in all cycles of software development
- Good master of English is required, knowledge of French is a plus



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Team spirit and initiative, ability work autonomously
- Knowledge of adversarial generative networks, Docker, HPC, or clinical data processing is a plus

Langues requises Français (Elémentaire) , Anglais (Avancé)

Localisation du stage Lausanne

Salaire mensuel 1750 CHF

Autres bénéfices et/ou compensations 2 jours de vacances / mois

Fichier (pdf) [Annonce Portail EPFL - Données synthétiques.pdf](#)

Remarque

Programmes Master

Masters concernés Mathématiques appliquées, Science et ingénierie computationnelles, Informatique, Systèmes de communication, Science des données