Feiyi Wang

Feiyi.wang@helsinki.fi   +1 (508) 873-7966
www.linkedin.com/in/feiyi-wang-fey   www.github.com/feiyiwang
www.orcid.org/0000-0002-8179-488X

EDUCATION

University of Helsinki, Helsinki, Finland   04/2021 - present
University of Edinburgh, Edinburgh, UK
Doctoral of Philosophy in Clinical Research
Concentrations: disease correlation, familial aggregation, outcome prediction, heritability
Dissertation: Risk factors in diabetes mellitus and autoimmune diseases using epidemiological methods, machine learning algorithms and multi-source healthcare data

Worcester Polytechnic Institute, Worcester, MA, USA   01/2015 - 05/2017
Master of Science in Data Science

Shanghai University of Finance & Economics, Shanghai, China   09/2008 - 06/2012
Bachelor of Arts in Economics Journalism

CAPABILITY

Healthcare Database: FinnGen, FinRegistry, SCI-Diabetes, UK Biobank, IQVIA, Symphony
Programing Languages: Python, R, SQL, HTML, CSS, JavaScript (JS)
Visualization Skills: Python - Django, Flask; R – Shiny; JS - D3, React, jQuery

EXPERIENCE

Data Scientist,   07/2020 – 12/2021
Institute for Molecular Medicine Finland, Helsinki, Finland
- Investigated significant increase of disease prevalence in 1998 in Finland introduced by outpatient data using Poisson regression, SVM and nationwide Finnish registries.
- Developed a healthcare-related question-answering platform using deep learning models (LSTM and pre-trained GPT) and FinRegistry

Oncology Commercial Data Scientist,   01/2017 - 06/2020
Takeda Pharmaceuticals, Cambridge, MA, USA
- Brand insights analytics and data visualization using Symphony specialty pharmacy data and IQVIA claims data and lab values
- Market insights and patient journey analyses (e.g., medication adherence, dose pattern classification, line of therapy, line prediction)
- Extracted cardio-toxicity attributes of kinases from online databases and built a searchable knowledge library for potential early detection of kinase inhibitors for clinical trial

Clinical Natural Language Processing Researcher,   01/2017 - 08/2018
Dana Farber Cancer Institute, Boston, MA, USA
- Extracted systematic treatment pathways (e.g., symptom, illness history, Gleason score, prostate-specific antigen, medication) from electronic health records (EHR) of prostate cancer patients using regular expressions and nltk package
- Established classification models to sectionize unstructured EHR with F1 score 0.96 and applied NLP to extract clinical events such as staging, code status, metastasis
- Developed a clinical analytics platform using Flask web framework

PUBLICATION

- Viippola E., …, **Wang F.**, …, Ganna A., FinnGen (2023). Data resource profile: nationwide registry data for high-throughput epidemiology and machine learning (FinRegistry), *International Journal of Epidemiology*

ORAL PRESENTATION

- Impact of parental autoimmune diseases on type 1 diabetes in offspring can be partially explained by genetic variants from HLA and non-HLA regions: a nationwide registry and biobank study. The American Society of Human Genetics, Washington, D.C., USA, November, 2023

- Can we improve multiple sclerosis risk stratification in optic neuritis patients presenting with clinically isolated syndromes? Novel application of a multiple sclerosis genetic risk score. UK Neuro-Ophthalmology Society, London, UK, March, 2023

POSTER PRESENTATION
- Familial aggregation of autoimmune diseases and type 1 diabetes in offspring. The European Society of Human Genetics, Glasgow, UK, June, 2023
- A smart tool for querying registry data. AI4Health Winter School, online, January, 2022

AWARD AND GRANT
University of Helsinki travel grant 2023, €2300
University of Helsinki travel grant 2022, €700

WORK LEAVE
Maternity leave, 01/2021 – 03/2021