Computational literature-based drug discovery

Dr. Anna Koroleva
Research associate,
korl@zhaw.ch

Dr. Manuel Gil
Lecturer, giln@zhaw.ch

Research project
Computational literature-based natural product drug discovery

Lead:
Dr. Manuel Gil, Fachstelle Computational Genomics

Duration:
Start: August 2019,
Financing for 24 months

Partner:
Dr. Manuel Gil, Principle Investigator; Dr. Anna Korolyova, Computational Linguistics; Dr. Andreas Lardos, Domain Expert Natural Product Research; Dr. Maria Anisimova, Computational Biology and Data Integration; Lirui Zhang, MSc Student Applied Computational Life Sciences; Dr. Evelyn Wolfram, Industrial Contacts Natural Product Research

Funding:
Grant from Health@N

Fachstelle Applied Computational Genomics

A team of natural product researchers and computational scientists funded by Health@N is developing a system for the automated discovery of new hypotheses and drugs from existing biomedical literature. To this end, neural networks and deep learning are used to integrate a large body of full texts with biomedical ontologies. The resulting knowledge base overcomes limitations of current systems.

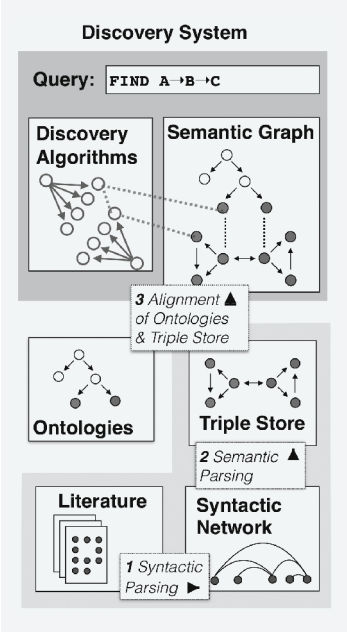
Natural products have successfully served as starting points for the discovery of important drugs. In 2015, Tu Youyou was awarded the Nobel prize in medicine for her discovery of artemisinin isolated from a natural product as a treatment against malaria, saving millions of lives. The compound was found through systematic analysis of ethnopharmacological texts. In pioneering work in the 1980s, Don Swanson found hidden links between pieces of knowledge in the scientific literature. With a manual algorithm, he found, for example, a paper showing that fish oil (A) can reduce vascular reactivity (B), and a different paper showing that a reduction in vascular reactivity can treat Raynaud's syndrome (C). Such knowledge can be linked: from A→B and B→C one can deduce A→C. Swanson hypothesized the novel link, which was later confirmed through clinical trials.

Algorithms operate on knowledge bases

Discovery of new knowledge from existing literature is nowadays referred to as Literature Based Discovery (LBD). Discovery algorithms commonly operate on knowledge bases. They provide formal representations of natural language texts. A common form are triples, semantic atomic data entities of the form <subject, predicate, object. An informal example would be <fish_oil, reduces, vascular_reactivity>. To build a knowledge base, the entities and relations are extracted from texts through controlled vocabularies, defining the scope of possible discoveries.

New applications emerge

In an interdisciplinary project between natural products research (ICBT) and computational science (IAS) funded by a Health@N grant, we will devise a new automated LBD system for natural product drugs. Our goal is to address limitations of state of the art systems. First, we will leverage full texts of biomedical publications, instead of titles and abstracts only. Second, LBD arose in the medical domain so the majority of use cases concern relations between diseases and potential drugs. Now new applications are emerging, where LBD is used to discover interactions involving genes, proteins, cells, receptors, biological processes, etc. The biomedical and bioinformatics communities have designed over 100 domain-specific interoperable ontologies for annotation and database integration. We will use relevant ontologies to add types of entities and relations that are of interest for LBD but are not covered in current systems. Third, current systems rely on simple dictionary look-up and rules for biomedical entity and relation extraction. For such complex tasks these methods are generally outperformed by natural language processing methods based on neural networks and deep learning. They are our methods of choice.



The system operates on a large body of biomedical literature, converted to a formal representation (steps 1 and 2). Ontologies are aligned to provide a common vocabulary (step 3). The resulting semantic graph can be explored via discovery algorithms

A video about our project, featured on the Health@N project website, has attracted the attention of the organisers (collaborators of Swanson) of the First International Workshop on Literature-Based Discovery. An article on our on-going work has been accepted for the peer-reviewed proceedings and presentation at the conference. ■

PiaBreed – Wearable Technology and AI in Veterinary Medicine

Prof. Dr. Thomas Ott, Head of research group Bio-Inspired Modeling & Learning Systems

AI (artificial intelligence) enabled wearable devices have become a major driver in developments in the human healthcare industry. This trend is now also being observed in the field of domestic animal and livestock breeding. However, there are special challenges associated with this application that result from a reduction in direct feedback from the patients. The Swiss start-up company Piavita is a pioneer in developing wearable technology for applications in veterinary medicine. Their sensors for equine patients can non-invasively measure a full set of vital signs. In collaboration with Piavita and the Swiss Institute of Equine Medicine at the University of Bern, we aim to further enrich and improve the functionality of the Piavita system. Our new Innosuisse funded project aims to develop AI algorithms for monitoring and predicting the birth of foals and equine ovulation. Both of these aims meet a clear need to increase process reliability and reduce costs in the multi-million dollar horse breeding market. ■

Wearable device for horses, ©Source Piavita

Neue Projekte

PiaBreed: Machine Learning zur automatisierten Ovulations- und Geburtsüberwachung am Pferd
Leitung: thomas.ott@zhaw.ch
Dauer: 01.01.2020 – 30.12.2022
Projektpartner: Universität Bern, Piavita AG, mitfinanziert durch Innosuisse

Weitere Projekte
zhaw.ch/ias/projekte

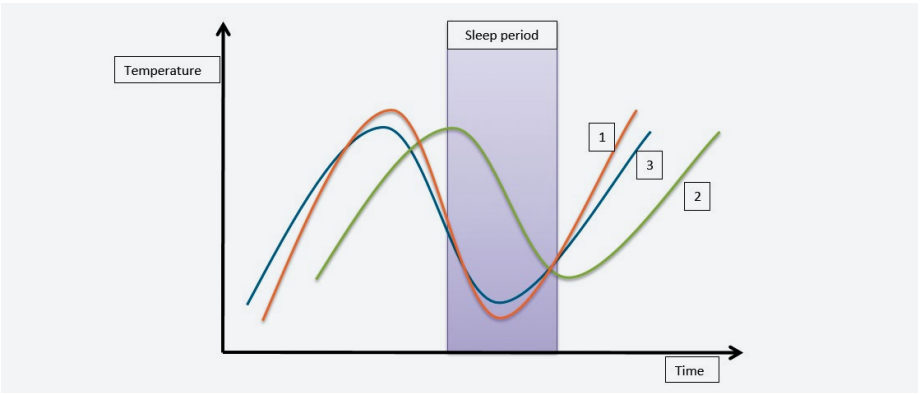
Weiterbildung

Diverse Kurse und Angebote
zhaw.ch/de/lsfm/institute-zentren/ias/weiterbildung/

Towards detection of neurodegenerative diseases using wearable devices and machine learning

Giovanni Schiboni, Research associate, scbo@zhaw.ch; **Krzysztof Kryszczuk,** Head of research group Predictive Analytics Group, krys@zhaw.ch

In late 2019, we started a new research project with the goal of developing a wearable device which could detect symptoms associated with Alzheimer's and Parkinson's Diseases (AD and PD). The device will detect fine differences in the circadian rhythm between healthy individuals and persons affected by neurodegenerative diseases, in "free living" conditions. The envisioned wearable device will incorporate physical sensors such as skin temperature and GreenTEG's proprietary heat flux sensors. We will use machine-learning techniques to estimate the core body temperature (CBT) trajectories, whose daily fluctuations reveal the circadian rhythm patterns. We are collecting a unique database of circadian rhythm recordings from both healthy individuals as well as patients affected by PD and AD. We are working on novel models to quantify the circadian rhythm shifts, which we expect to be indicative of neurodegenerative diseases (Figure). In the intended use mode, the developed wearable device will collect circadian rhythm data from the user, and the information extracted will be evaluated using the models. Similarity with the AD/PD models will trigger a warning of a possible onset of neurodegenerative disease and an early warning of possible health condition will be issued, prompting the wearer to consult a healthcare expert for a clinical diagnosis. ■



Models of PD (1), AD (2) and healthy (3) individuals; 1. PD patients (orange curve): have a less pronounced nadir (lowest CBT during the night) and lowered CBT. AD patients (green curve): phase-shift plus increased CBT but without fluctuations during nadir