

Evolutionäre Optimierung von Klassifiziererverbunden zur Vorhersage antimikrobieller Peptide

Sebastian Spänig

AG Bioinformatik (Prof. Dr. Dominik Heider)
sebastian.spaenig@uni-marburg.de

Zusammenfassung

Eine Strategie zur Bekämpfung dieser multiresistenten Krankheitserregern umfasst die Verwendung natürlich vorkommender, antimikrobiell wirksamer Peptide (AMP). In den letzten Jahren wurden deshalb Deskriptoren und maschinelle Lernverfahren zur automatisierten und genauen Klassifizierung von AMPs entwickelt [1].

Weiterhin erfordern Algorithmen des maschinellen Lernens eine feste Länge der Merkmalsvektoren sowie eine metrische Ausprägung dieser. So beinhaltet ein wesentlicher Teil der Klassifizierung die Konstruktion von repräsentativen Kodierungen der Peptidsequenz. Zu diesem Zweck wurden im Rahmen von Klassifizierungsstudien bereits eine Vielzahl an sequenz- und strukturbasierter Kodierungen entwickelt [2]. Weiterhin konnte bereits im Kontext von HIV gezeigt werden, dass mit verschiedenen kodierten Datensätzen trainierte Classifier Ensembles die Genauigkeit der Vorhersage steigern können [3].

Ziel dieser Arbeit soll daher die Entwicklung eines Frameworks zur evolutionären Optimierung (z.B. genetischen Algorithmen [4]) sog. Ensembles sein, mit dessen Hilfe die Vorhersage von AMPs verbessert werden kann. Zu diesem Zweck sollen im Rahmen der Abschlussarbeit 1) unterschiedliche Fusions- und Evolutionsstrategien umgesetzt sowie 2) die Software so gestaltet werden, dass eine einfache Erweiterbarkeit, bspw. eine Parallelisierung mit Apache Spark, gewährleistet ist.

Literatur

- [1] Mahlapuu, M., Håkansson, J., Ringstad, L., Björn, C. (2016). Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. *Front. Cell. Infect. Microbiol.* 6, 194.
- [2] Spänig S. and Heider D. (2019). Encodings and models for antimicrobial peptide classification for multi-resistant pathogens. *BioData Mining* 12,7.
- [3] Dybowski J.N. et al. (2011). Improved Bevirimat resistance prediction by combination of structural and sequence-based classifiers. *BioData Mining* 4,26.
- [4] Pyka M. et al. (2011). Dynamic causal modeling with genetic algorithms. *Journal of neuroscience methods.* 194(2), 402-406.

Bitte melden Sie sich bei Interesse per Email bei:
sebastian.spaenig@uni-marburg.de
Stand: 04.04.2019

