

## Bachelorarbeit in Bioinformatik

Prof. Dr. Dominik Heider  
dominik.heider@uni-marburg.de

### **Thema: Einfluss von sozioökonomischen Faktoren auf Biodiversität - Feature Generation und Machine Learning**

Natürlich vorkommende Ökosysteme sind komplexe Netzwerkstrukturen interagierender Organismen und ihrer unbelebten Umwelt. Das bedeutet, dass sich nicht nur Organismen untereinander beeinflussen (z.B. durch Prädation oder Symbiose), sondern die Lebensfähigkeit eines Organismus auch von dem Zustand der Umwelt abhängt und diese Umwelt von den Organismen geformt wird.

Eine zentrale Maßzahl für den Zustand eines Ökosystems ist dessen Biodiversität, d.h. die Vielfalt der Organismen in diesem Ökosystem. Ein starkes Abnehmen der Biodiversität eines Ökosystems kann als Indikator für eine Veränderung oder sogar einen Zusammenbruch des unterliegenden Ökosystems genutzt werden (1,2).

Die größten Biodiversitätsverluste der letzten Jahrzehnte gehen auf menschliche Einflüsse zurück. Darunter fallen unter anderem (bauliche) Veränderung der Lebensräume, Umweltverschmutzung und menschengemachter Klimawandel. Viele dieser Faktoren sind anhand von GPS-Daten, Online-Weltkarten und öffentlicher Datensätze nachzuvollziehen.

In Zusammenarbeit mit der Arbeitsgruppe Boenigk der Universität Duisburg-Essen haben wir Zugang zu einem Datensatz der mikrobiellen Biodiversität von mehr als 200 europäischen Seen (3). In diesem Projekt sollen zunächst auf der Basis von GPS-Daten Informationen zu sozioökonomischen Faktoren gesammelt werden, die einen Einfluss auf die Ökosysteme der Seen haben könnten (feature generation). Mit Hilfe von feature selection-Ansätzen wie PCA sollen dann relevante Variablen oder Variablengruppen gesucht werden. Schließlich soll mit Hilfe von Machine Learning-Modellen wie z.B. random forests, und nicht-linearen Regressionsmodellen festgestellt werden, welcher dieser features einen Einfluss auf die Biodiversität von Seen hat.

#### **Literatur**

[1] B. Worm, E. B. Barbier, N. Beaumont, J. E. Duffy, C. Folke, B. S. Halpern, J. B. C. Jackson, H. K. Lotze, F. Micheli, S. R. Palumbi, E. Sala, K. A. Selkoe, J. J. Stachowicz & R. Watson (2006). Impacts of Biodiversity Loss on Ocean Ecosystem Services. *Science*, 314, 787-790.

[2] L. Grossmann, M. Jensen, D. Heider, S. Jost, E. Glücksman, H. Hartikainen, S. S. Mahamdallie, M. Gardner, D. Hoffmann, D. Bass, J. Boenigk: Protistan community analysis: key findings of a large-scale molecular sampling. *ISME J.* 2016, 10(9):2269-79.

[3] L. Grossmann, D. Beisser, C. Bock, A. Chatzinotas, M. Jensen, A. Preisfeld, R. Psenner, S. Rahmann, S. Wodniok & J. Boenigk (2016). Trade-off between taxon diversity and functional diversity in European lake ecosystems. *Molecular Ecology*, 25, 5876-5888.

**Bitte melden Sie sich bei Interesse per Email bei:**  
**Prof. Dr. Dominik Heider - dominik.heider@uni-marburg.de**

**20.04.17**