



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN



Prof. Dr. Christian L. Müller
Ludwig-Maximilians-Universität München
Institut für Statistik
Ludwigstr. 33
80539 München
christian.mueller@stat.uni-muenchen.de

M.Sc. Thesis Proposal: *Learning graphical models to explore the relationship between human gut microbiota and cardiovascular diseases.*

Objective:

Many diseases are multifactorial in origin, meaning that they are caused by a combination of genetic and environmental components. To date, a considerable number of genetic variants have been identified that are associated with almost every multifactorial disease. These independent genetic factors are often common, occurring frequently in the absence of disease, and therefore cannot yet be used to predict disease. The purpose of this M.Sc. thesis is **to explore the casual effect of those factors and cardiovascular disease using state-of-art machine learning and statistical methods**. We start with exploratory data analysis of the sample from *LifeLines* cohort and learning its graphical representation. Once we understand the structure of the graph (network), we can apply methods, e.g., Gaussian Graphical Models, to answer such questions as:

- “Is there a relationship between microbial ecosystems and cardiovascular diseases?”
- “If there is a relationship, what type of the relationship (e.g., a non-linear relationship)?”
- “How does microbial abundances change with respect to the type of a cardiovascular disease?”

Plan and deliverables:

LifeLines is a population cohort of over 165 000 participants from the northern Netherlands that covers multiple generations of participating families and focuses on determinants for multifactorial diseases, e.g., cardiovascular diseases. The cohort includes detailed information on phenotypic and environmental factors, as well as health status. Our dataset contains meta-genomic sequencing, 16S sequencing of stool samples (domain V4) and Second round of follow-up of 1,200 participants.



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN



During the thesis you will get a methodological support in your analysis and practical experience with best coding practices and frameworks in Python. Also, you will have a chance to work with a real-world data and to contribute to open-source project Qiime2 which are both good for your resume and practical skills development. A write-up in thesis form and commented code on GitHub are mandatory deliverables at the end of the thesis.