



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](mailto:christian.mueller@stat.uni-muenchen.de)

Master's thesis proposal: *Identifying chemical features predictive of bacterial pathogen stress response.*

**Background:** The prevalence of resistance to current antibiotic treatments is a major health concern that is set to increase in the coming years. Investigating how bacterial pathogens respond to chemical treatments (such as antibiotics), is key in order to develop targeted antimicrobial therapies. With this goal in mind, our *StressRegNet* [1] consortium has generated a dataset, where the expression of key survival genes in *Salmonella* and *Campylobacteri jejuni* is measured in response to a large chemical library.

**Objective:** The purpose of this project is to develop a strategy for identifying chemical features that are predictive of a given outcome. Specifically, the student will work with the data generated by the *StressRegNet* consortium to identify the chemical features that best explain the observed changes in gene expression. Depending on the student's interest, and how the project develops, various methods from the field of high-dimensional statistics (e.g. LASSO regression, clustering, deep-learning) will be used for this purpose. While this is interdisciplinary, no knowledge of biology or chemistry is necessary. Collaboration with domain experts will be key, with the potential to experimentally validate findings from the statistical analysis.

**Plan and deliverables:** The development of a reproducible pipeline that receives molecular structures and can identify chemical features that are predictive of a given outcome. This pipeline should be able to analyze the data from *StressRegNet* but should also be applicable to other datasets. A typical result from this analysis can be "Chemical feature X is predictive of an increased expression of gene A. Therefore, compounds with feature X are likely to cause increased expression of gene A." A willingness to collaborate with domain experts is key. A write-up in thesis form and commented code on GitHub are mandatory deliverables at the end of the thesis.



## References:

1. <https://bayresq.net/en/projekte-stressregnet-en/>