

AI for Biotechnology

Exercise 3

Prof. Dr. Dominik Grimm

Bioinformatics Research Lab

TUM Campus Straubing for Biotechnology and Sustainability

Exercise E3.1

We have developed a novel diagnostic test in our newly founded startup. This test screens for genetic mutations to then predict if a certain patient suffers from a certain rare disease. The following measures are recorded by our scientific team in a clinical trial study:

	Disease	Healthy	Total
Prediction Disease	190	210	400
Prediction Healthy	10	3590	3600
Total	200	3800	4000

- What is the accuracy, precision, recall, F1-score, specificity, false positive rate and the Matthew's Correlation Coefficient? (Compute the numbers of paper)
- The PPV (precision) is defined as $\frac{TP}{TP+FP}$. Think about how you would compute the Negative Predictive Value (NPV) and compute this value (on paper).
- Interpret your results. Is the developed diagnosis test a good test? What are the strength and weaknesses?

Exercise E3.2

Download the Jupyter Notebook `Exercise3.ipynb` and solve the exercises in this Jupyter Notebook.