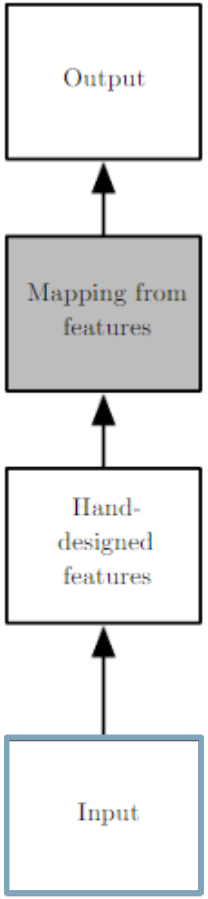
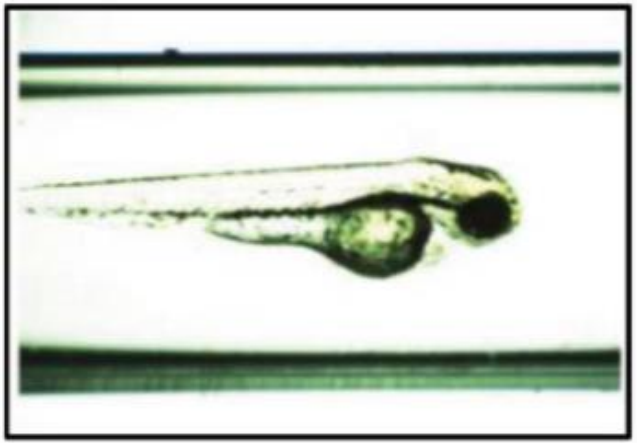
General concept of data science

Classic machine learning

Input
Data acquisition
Data annotation
Data preprocessing
Data augmentation

Hand-designed features
Mapping from features
Output





Mark Schutera, Thomas Dickmeis, Marina Mione, Ravindra Peravali, Daniel Marcato, Markus Reischl, Ralf Mikut, and Christian Pylatiuk. Automated phenotype pattern recognition of zebrafish for high-throughput screening. *Bioengineered*, 7(4):261–265, 2016.

Mark.Schutera@mailbox.org

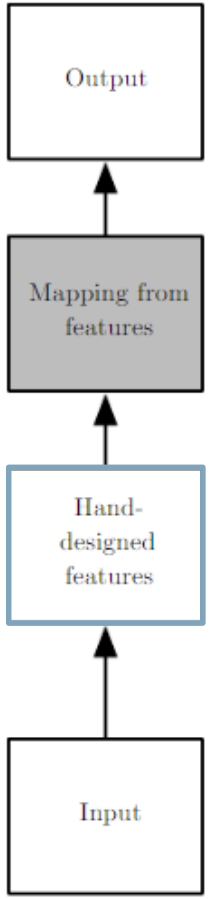
General concept of data science

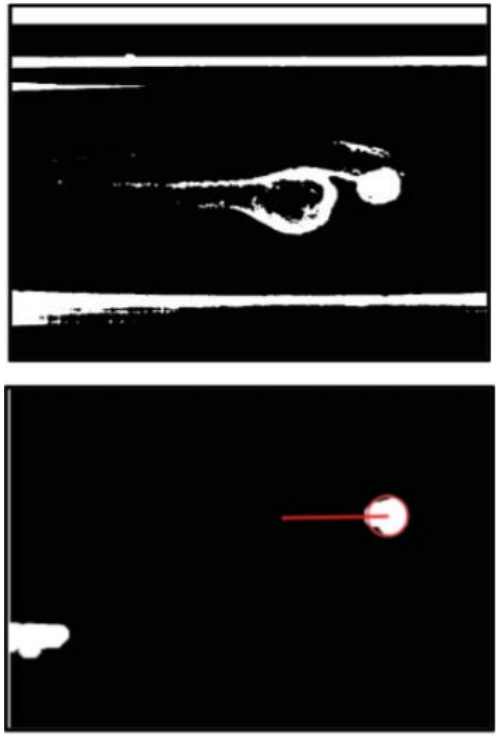
Classic machine learning

Input

Hand-designed features
Domain knowledge
Trial and error
Feature selection methods
Feature set selection methods

Mapping from features
Output





General concept of data science

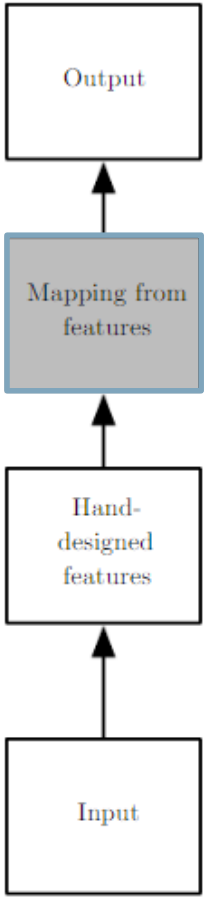
Classic machine learning

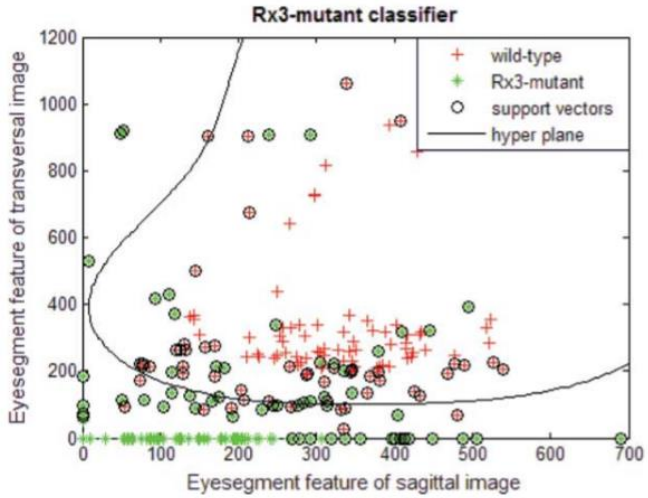
Input
Hand-designed features

Mapping from features
Learn pattern from data..
..Support vector machine
..Random forest

..Unsupervised – Class label is not known
..Supervised – Class label is known

Output



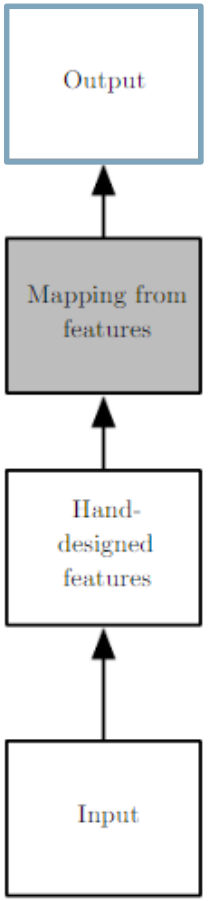


General concept of data science

Classic machine learning

Input
Hand-designed features
Mapping from features

Output
Inference step on data not seen during training

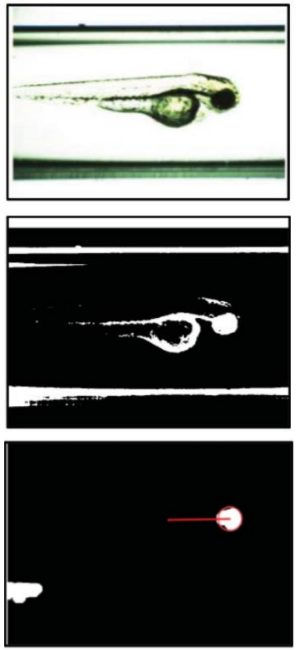


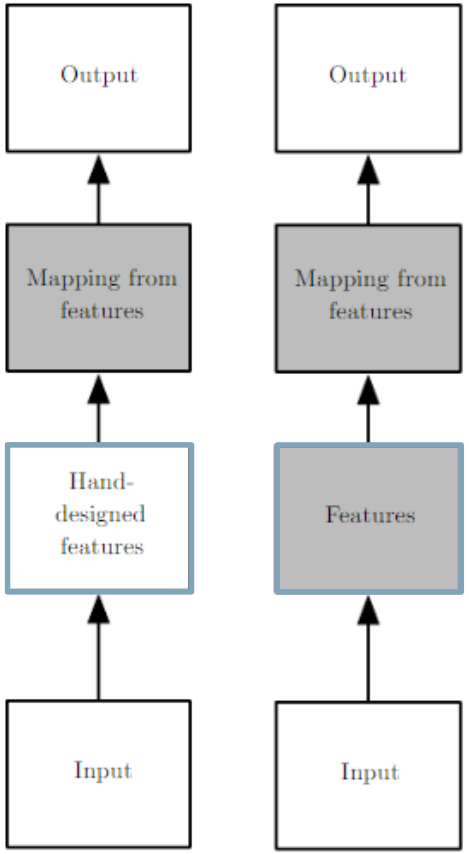
0 wildtype

1 mutant

Classic machine learning to deep learning

Classic machine learning
Hand-designed features





Representation learning
Data-driven feature selection