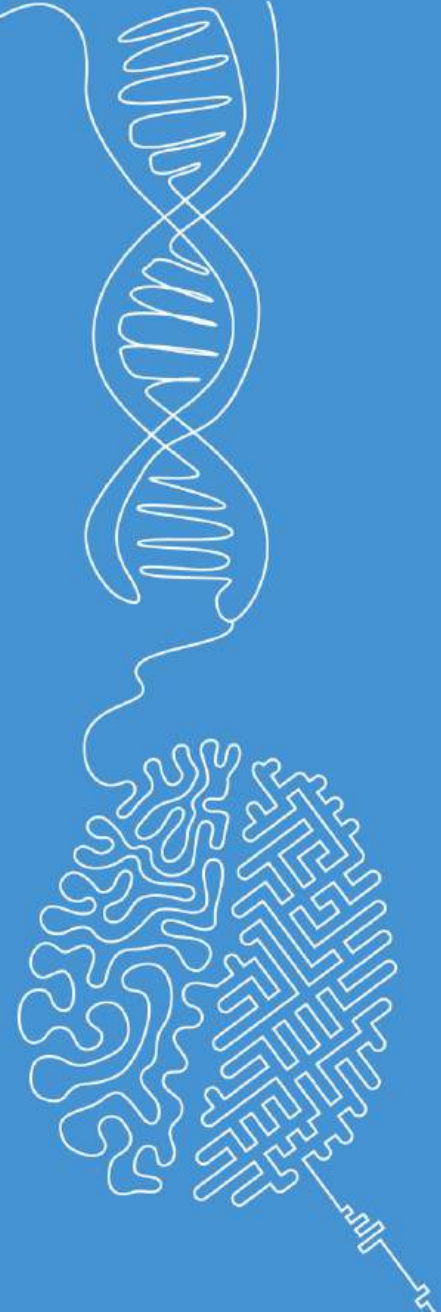


Dataset proposals

Machine Learning

Norman Juchler



Goal of the project

- Implement an entire machine learning pipeline
 - Problem formulation and data description
 - Exploratory data analysis and data visualization
 - Method selection and model training
 - Model testing and performance analysis
 - Discussion
- Types of problems
 - Supervised learning
 - Unsupervised learning
 - Dimensionality reduction

Rules for the project work

- Each student works on a different dataset!
- The project is implemented entirely in a single Jupyter notebook
- The submission consists of a ZIP folder containing
 - a functional Jupyter notebook
 - a PDF version of the notebook with all its output
 - data and other resources used
- **Note:** The size of the submission (including data) must not exceed 50MB
- The project must be individual work
 - Copying an existing notebook is not permitted.
 - You may have to answer individual questions about your project.

Hints

- Choose a “relevant” problem
- Choose a “solvable” problem
- Ideas to find public datasets:
 - [Kaggle](#)
 - [Google dataset search](#)
 - [Open data portal](#) (Bundesamt für Statistik)
 - [List of datasets for ML](#) (Wikipedia)
 - ...

Proposals

NAME OF STUDENT

- **Dataset / project name:** PLEASE ADD
- **Brief description:** PLEASE ADD
- **Type of problem:** PLEASE ADD
- **Type of data:** PLEASE ADD
- **Dataset in numbers:** PLEASE ADD
- **Online source:** PLEASE ADD

**PLEASE ADD
REPRESENTATIVE SAMPLES
(OR SCREENSHOTS IF FILE
SIZE IS LARGE OR 3D)**

Nebuchadnezzar Juklerassar

- **Dataset / project name:** Distinguish between dangerous and benign aneurysms based on their shape
- **Brief description:** Aneurysms are pathologic dilations of (cerebral) arteries that may rupture and cause a bleeding in the brain. Being able to predict the rupturing of an aneurysm could save lives!
- **Type of problem:** Dimensionality reduction and supervised learning
- **Type of data:** Features X: numerical descriptions of the aneurysm shape. Target labels y: Rupture status
- **Dataset in numbers:** 750 aneurysms, 170 precomputed features
- **Online source:** <https://zenodo.org/records/6678442>

source	dataset	curvature	curvature	curvature	c
hug2016	p043_HAARCEcDAAQDQcbHgANDRQM	2.10899	0.214448	3.36637	
hug2016	p044_BBMDfXESDBMcEwcVBhMBExQC	1.31256	-0.137086	2.39603	
hug2016	p046_FwQADBEggwQBGwMdHwAADBAK_RICA	3.67321	-1.06886	6.13175	
hug2016	p091_ABMXCAM8FT4eAaABwUQFAR_LICA	8.84854	-0.717287	13.2938	
hug2016	p092_ABMXCAM8FT4eAaABwUQFAR_RICA	2.61218	0.383217	4.09299	
hug2016	p094_ERUeAB4WBgDSCwcbHgAABAEU	3.53995	-0.973234	5.03718	
hug2016	p095_EhMHDx4NAB8BEAg8FMMFwAW_LICA_1	2.206	0.444103	4.77592	
hug2016	p097_Eg4HEwEdGw8NChMdAxQHCxcX	1.77812	0.741736	2.94288	
hug2016	p099_FxMbBAMBHAYWFxgdGwQdDBsf	3.51399	0.210701	4.97989	
hug2016	p101_EwAeDgoWEw0cHwUOHg4UBhQU	3.77949	-0.897367	6.39671	
hug2016	p103_BAgTDgUBAQAKABgGHw4bFgYZ	2.37984	0.0972338	4.28309	
hug2016	p104_FhMbBBQNHB				
hug2016	p109_ARQbDwQIHAA				
hug2016	p114_FxQVDRUaDBC				
hug2016	p115_EgAeDBUaDAa				
hug2016	p117_BA4IAAMBDA4e				
hug2016	p119_AgACCAMJAAU				
hug2016	p120_HA4AB84SHR8				
hug2016	p121_FwgDFBUEDAg				
hug2016	p122_BAQWBAMLHR				
hug2016	p124_FAQeFBMJDAQ				
hug2016	p126_FAQBCCQuBAQC				
hug2016	p126_FAQBCCQuBAQC				
hug2016	p127_AwlaFhEaCAwS				
hug2016	p128_EggcChUaBh8a				
hug2016	p128_EggcChUaBh8a				
hug2016	p129_ERUeAB4WBgD				
hug2016	p131_EwglABQHAD4f				
hug2016	p133_AxUADhgAFwoV				
hug2016	p134_EwkTEwNAT4D				
hug2016	p135_FAQtEh8dAQAR				
hug2016	p136_Fw4fBAoVHQa				

source	dataset	hospital	status	location	side	sex
hug2016	p043_HAARC Geneva		unruptured	ICA oph	left	female
hug2016	p044_BBMDf Geneva		unruptured	VA V4	left	female
hug2016	p046_FwQAC Geneva		unruptured	MCA bif	right	male
hug2016	p091_ABMXC Geneva		ruptured	ICA pcom	left	female
hug2016	p092_ABMXC Geneva		unruptured	ICA oph	right	female
hug2016	p094_ERUeA Geneva		unruptured	ICA cav	right	female
hug2016	p095_EhMHC Geneva		unruptured	MCA bif	left	female
hug2016	p097_Eg4HE Geneva		unruptured	ICA oph	right	female
hug2016	p099_FxMbB Geneva		unruptured	ICA oph	right	female
hug2016	p101_EwAeD Geneva		unruptured	ICA cav	left	female
hug2016	p103_BAgTD Geneva		unruptured	ICA oph	left	female
hug2016	p104_FhMbB Geneva		unruptured	BATip	midline	female
hug2016	p109_ARQbC Geneva		unruptured	ICA oph	right	female
hug2016	p114_FxQVD Geneva		unruptured	ACAA1	right	female
hug2016	p115_EgAeD Geneva		unruptured	ICA oph	right	female
hug2016	p117_BA4IAA Geneva		unruptured	ICA oph	left	female
hug2016	p119_AgACC Geneva		unruptured	AComA	midline	female
hug2016	p120_HA4AB Geneva		unruptured	ICA bif	right	female
hug2016	p121_FwgDF Geneva		unruptured	ICA oph	left	male
hug2016	p122_BAQWI Geneva		unruptured	MCA M1	left	female
hug2016	p124_FAQeF Geneva		unruptured	ICA oph	right	female
hug2016	p126_FAQBCC Geneva		unruptured	MCA bif	left	female
hug2016	p126_FAQBCC Geneva		unruptured	ICA pcom	left	female
hug2016	p127_AwlaFr Geneva		unruptured	ICA oph	left	female
hug2016	p128_EggcCl Geneva		unruptured	ICA para	left	female
hug2016	p128_EggcCl Geneva		ruptured	AComA	midline	female
hug2016	p129_ERUeA Geneva		unruptured	ICA cav	left	female
hug2016	p131_EwglAE Geneva		unruptured	BATip	midline	female

Shape features (top) and clinical parameters about the cases

