

Data Stewardship DMP Exercise

ID (HTTP-ORCID): <https://doi.org/10.5281/zenodo.2648751>

Created: 22-04-2019

Modified: 22-04-2019

Language: en

Description:

	Data Officer	Moritz Leidinger, moritz.leidinger@student.tuwien.ac.at https://orcid.org/0000-0002-4537-6648 (HTTP-ORCID)
I	Data Characteristics	
I.1	Description of the data	<u>UCI Arrhythmia Dataset</u> This experiment aims to apply two machine learning models to solve a classification task for two different datasets. Each dataset is evaluated with each of the machine learning models, using different parameter settings and preprocessing strategies to compare the respective results and analyze across datasets and/or machine learning models. Type: dataset Language: en Keywords: Breast Cancer, Arrhythmia
II	Documentation and Metadata	
II.1	Metadata standards	
II.2	Documentation of data	Issued:
II.3	Data quality control	<ul style="list-style-type: none">• We use a readme file to describe the data quality and provide the source code of the experiment.
III	Data availability and storage	
III.1	Data sharing strategy	ID-1: https://doi.org/10.5281/zenodo.202648713 (HTTP-ORCID) <u>UCI Repository</u> : A publicly available dataset about Arrhythmia. Available till: Size: <10MiB Data access: open URL: https://archive.ics.uci.edu/ml/machine-learning-databases/arrhythmia/arrhythmia.data Format: text/csv

III.2	Data storage strategy	Cost(s) <u>UCI Repository</u> <ul style="list-style-type: none"> • UCI Repository • • • • • US • • unknown
IV	Legal and ethical aspects	
IV.1	Legal aspects	License: <ul style="list-style-type: none"> • Valid from: 2019-04-22 https://opensource.org/licenses/MIT
IV.2	Ethical aspects	unknown Must be preserved to keep research & development between stakeholders intact. unknown Ethical issues: <ul style="list-style-type: none"> • • unknown •