Dataset allocation

Group	Dataset
Nils Kulmbacher, Hubert Perlinski	ADULT
Albin Lindqvist, Helena Mūrniece	MUSHROOM
Pieter Joep Duijvestein, Tristan van Weeren	MUSIC
Maarten Hoogeboom, Roel Lust	CHURN
Bart ten Brinke, Melis Çevik	LOAN
Yujie Ma, Xinyu Wang	HEART DISEASE
Lotte Bulder	MUSHROOM
Xinhui Wang, Yunzhuo Wang	ADULT
Kabir Agarwal, Canberk Döven	CHURN
Ionel Popescu, Piotr Zieliński	BANK MARKETING
Constantin Yogeshwar, Patra Haralabus	MUSIC
Xueni Chen, Jingxian Mei	HEART DISEASE
Neil Basson, Sam Heinz	BANK MARKETING

Dataset info

Dataset	Source	Link	# samples	# features	target name	target meaning	Notes
ADULT	UCI ML repository	https://archive.ics.uci.edu/dataset/2/adult	47621	14	income	Predict whether income is <=50K or >50K	
MUSHROOM	UCI ML repository	https://archive.ics.uci.edu/dataset/73/mushroom	5644	22	poisonous	Predict whether mushroom is poisonous or not (p—poisonous; e—eatable)	
LOAN	Kaggle	https://www.kaggle.com/datasets/burak3ergun/loan-data-set/data	480	12	Loan_Status	predict whether eligible for loan (yes/no)	
BANK MARKETING	UCI ML repository	https://archive.ics.uci.edu/dataset/222/bank+marketing	7842	16	y	predict if the client will subscribe a term deposit (Y—yes/N—no)	
HEART DISEASE	UCI ML repository	https://archive.ics.uci.edu/dataset/45/heart+disease	297	13	num	predict disease presence (1) vs absence (0)	The original target has integer values 0-4, where 0=absence and values 1-4 are different levels of presence. As stated in the dataset documentation, most published work focuses on absence (0) vs. Presence (values 1-4). In the dataset I provide I have already recoded the target into 0/1
MUSIC	Kaggle	https://www.kaggle.com/datasets/solomonameh/spotify-music-dataset/data	4829	29	popularity	predict popularity (high/low)	
CHURN	Kaggle	https://www.kaggle.com/datasets/kapturovalexander/customers-churned-in-telecom-services/data	7032	29	Churn	predict whether the customer has churned (Yes/No)	