orange
DATA MINING

Data Mining
Fruitful and Fun

Open source machine learning and data visualization.

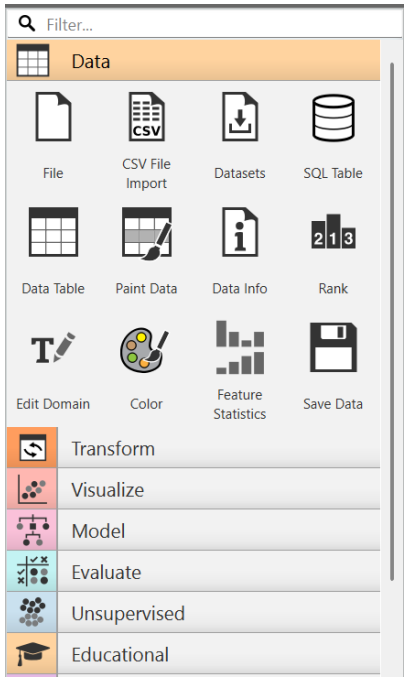
Simple AI Methods for Demetia: Orange

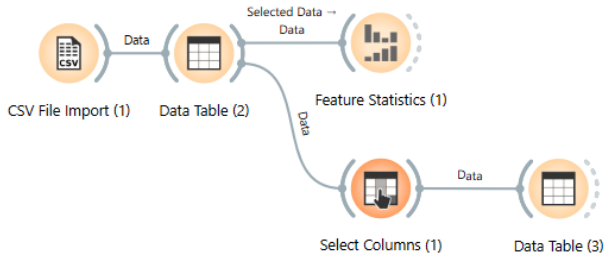
SAMUEL MADDOX, SCHOOL OF COMPUTING SCIENCES

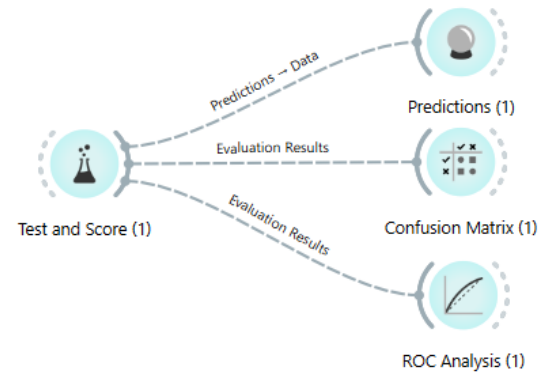
HTTPS://ORANGEDATAMINING.COM/

Orange: A Tool for Visual Data Science

- Open-source, easy installation
- Drag-and-drop widgets
- Enables AI without coding

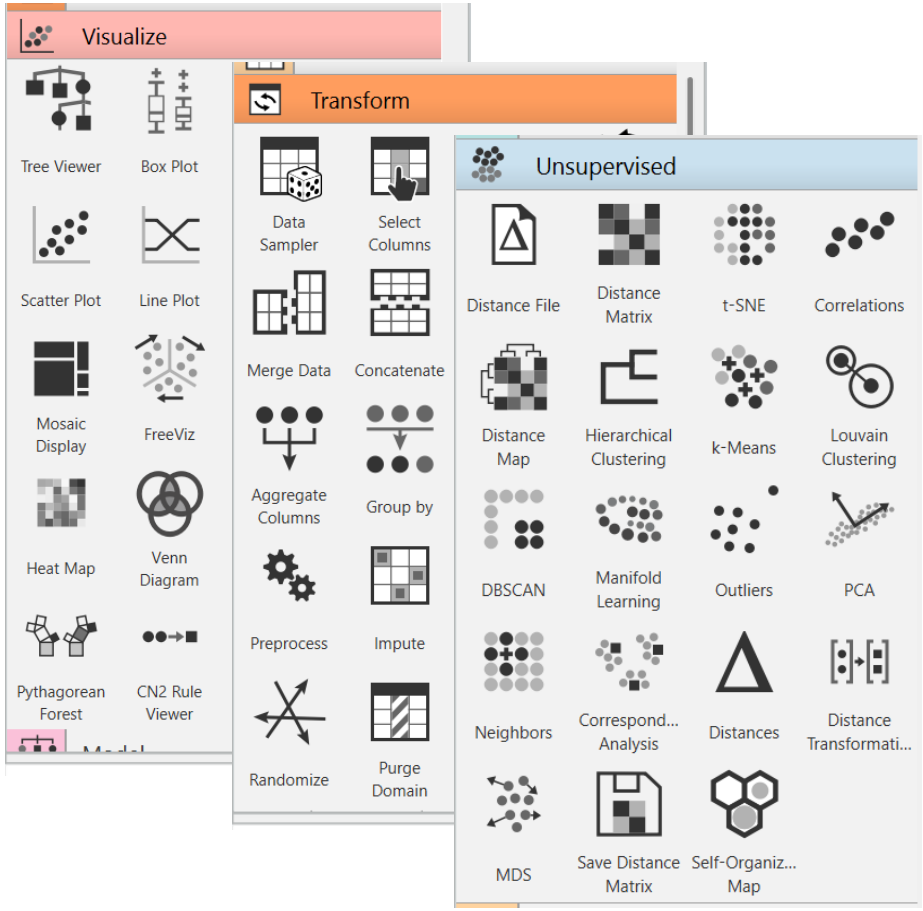






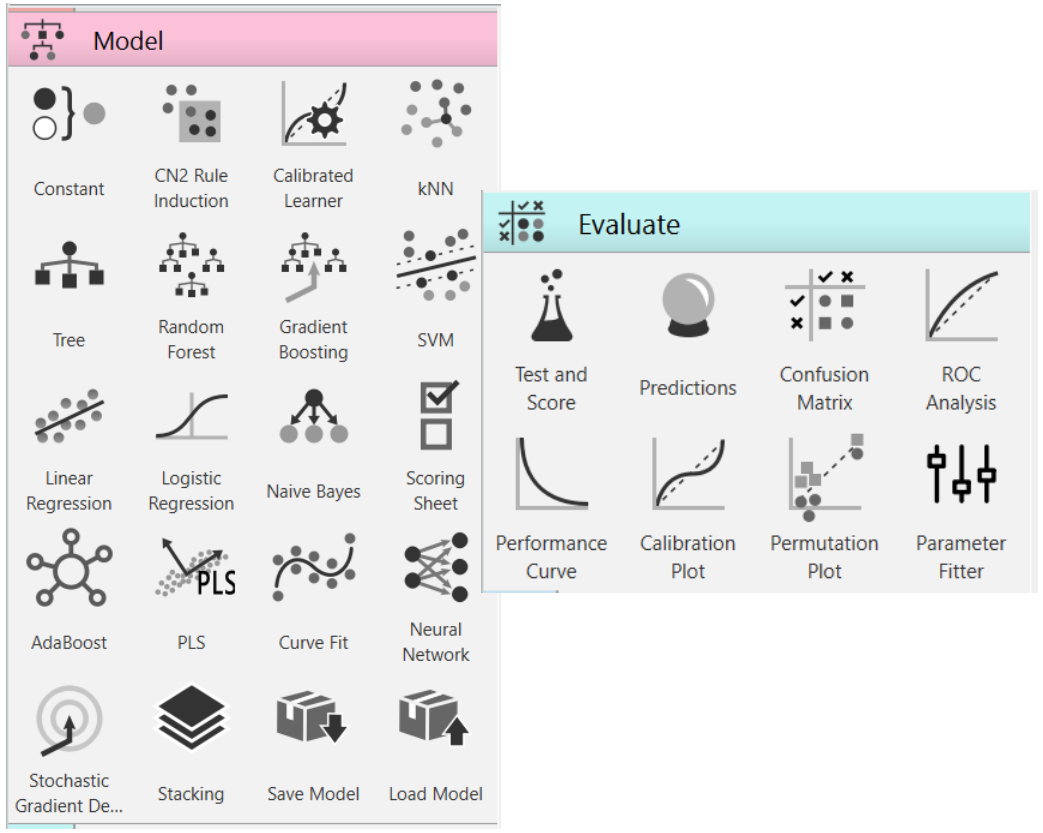
Capabilities for Healthcare Data Analysis

- Easy data loading & preparation
- Interactive data exploration
- Essential data analysis tools



Applying Machine Learning Techniques

- Access ML algorithms visually
- Build predictive models easily
- Train, test, and evaluate

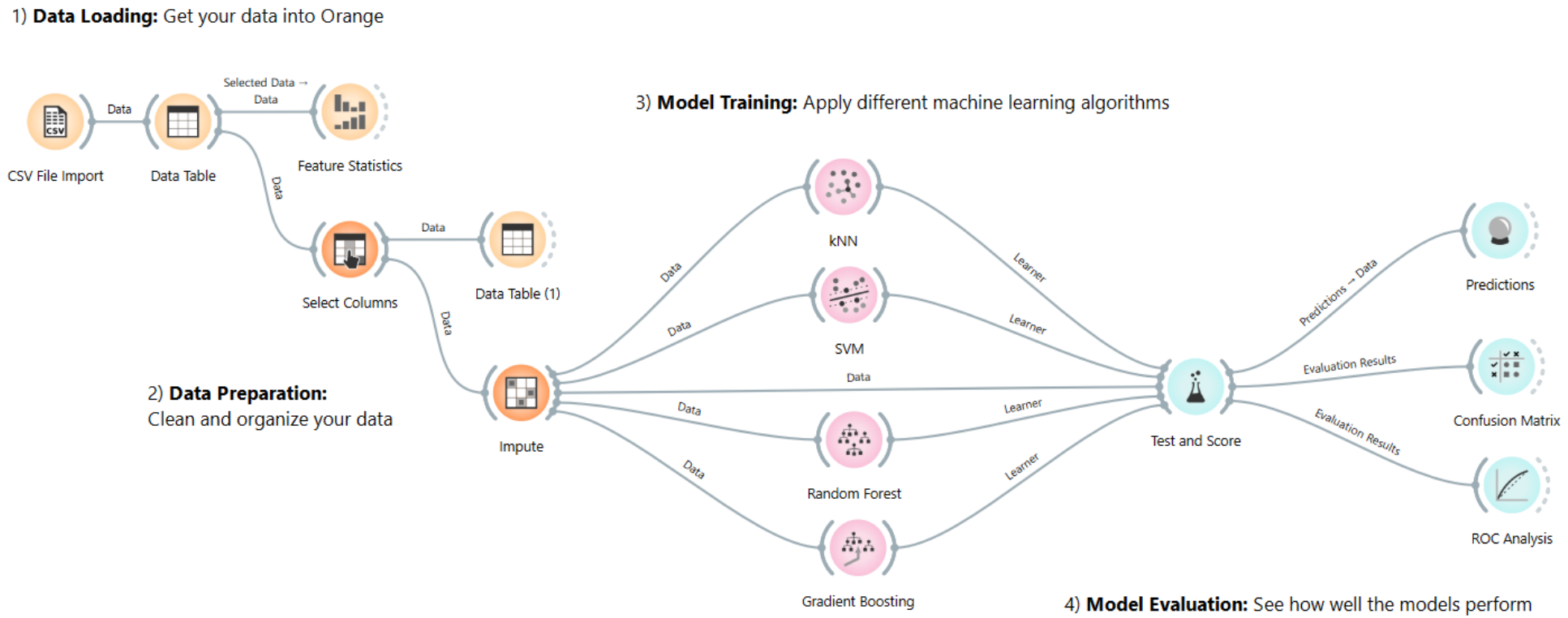


Case Study: ADNI Group Classification

- Applying Orange to ADNI data
- Goal: Classify patient groups
- Explore various model performance metrics

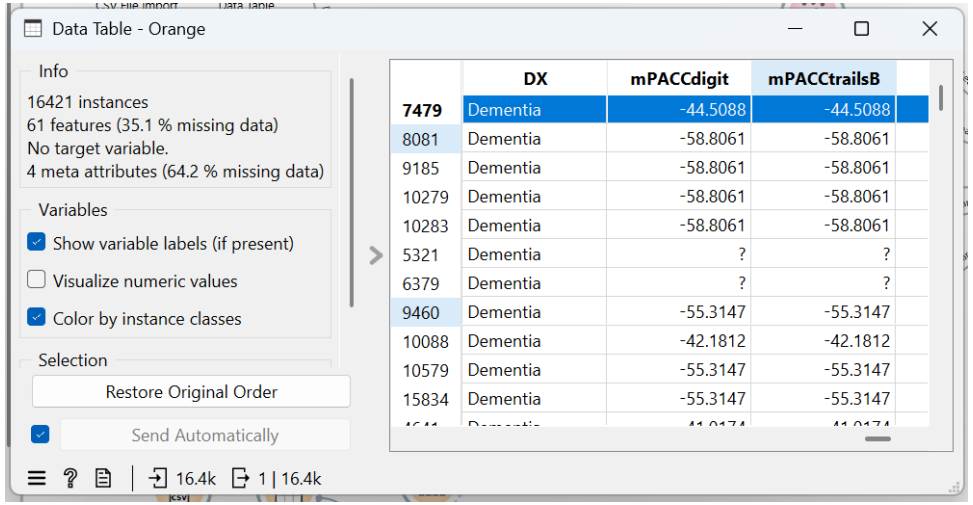
ADNI Workflow in Orange

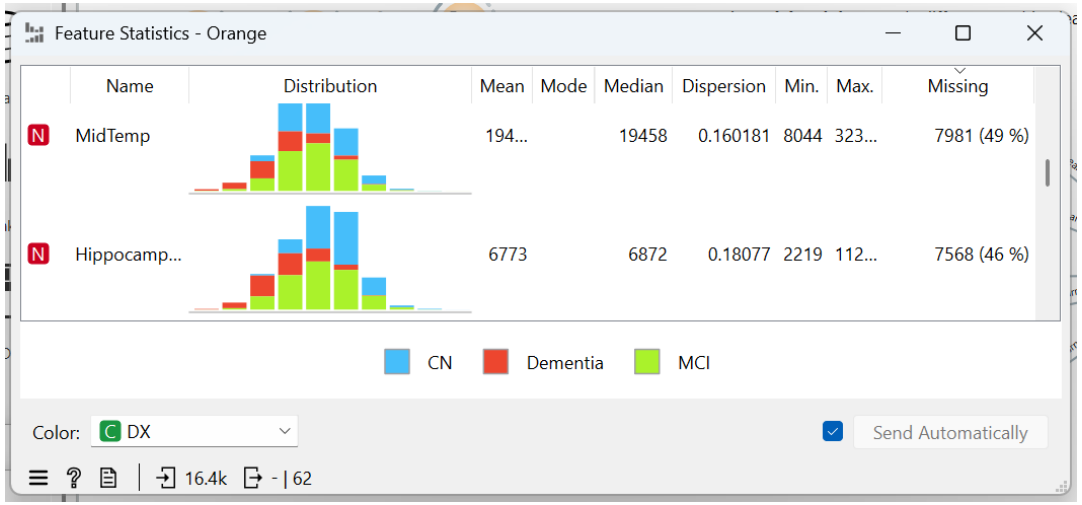
- Connect widgets to build the analysis pipeline



ADNI Workflow in Orange

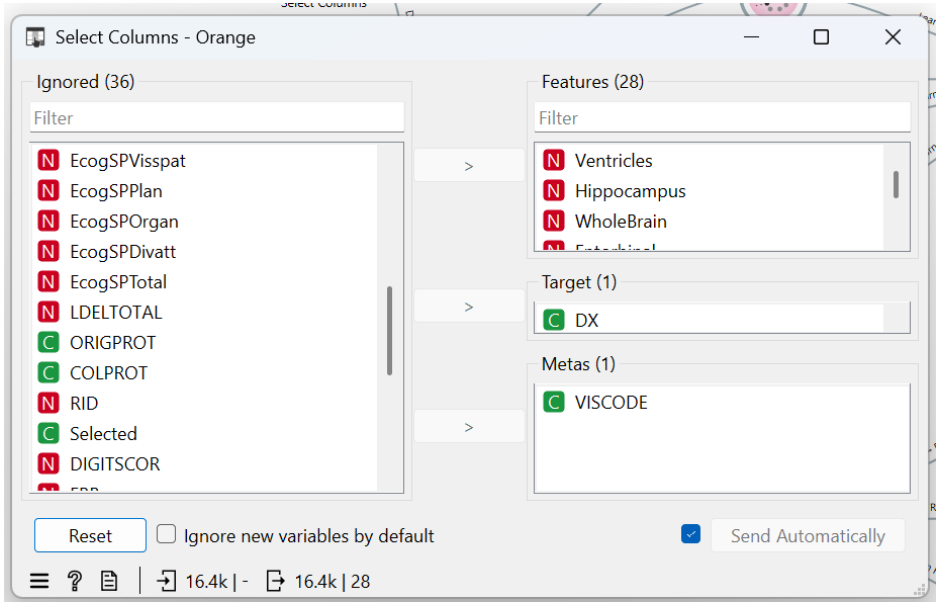
- 1 – Data Loading:

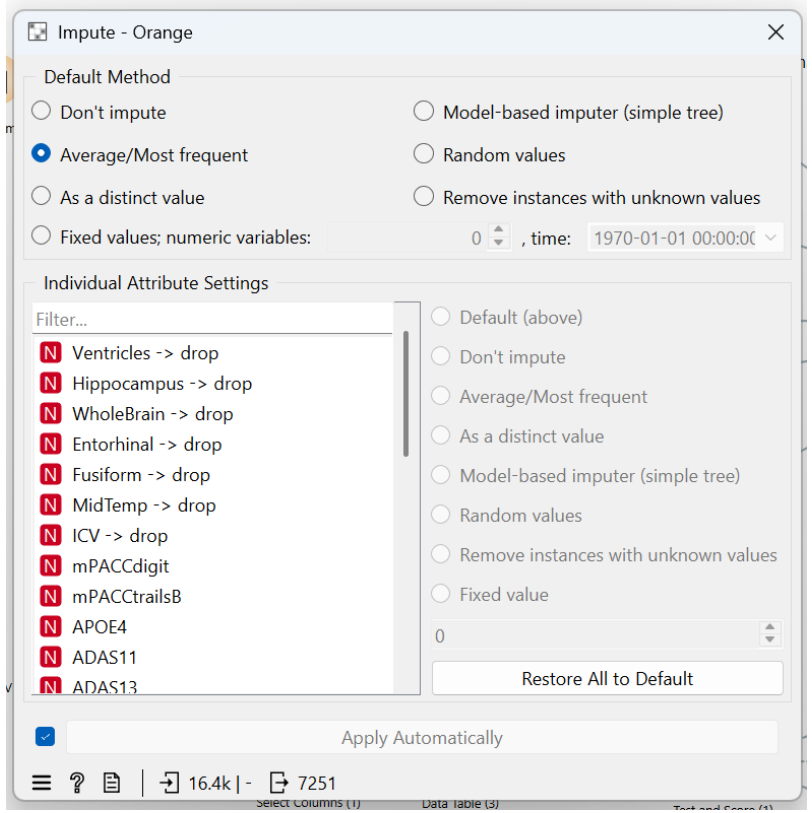




ADNI Workflow in Orange

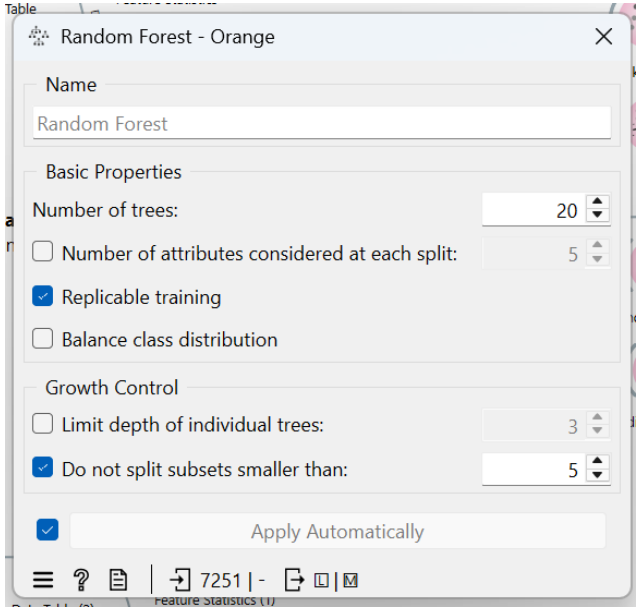
- 2 – Data Preparation:

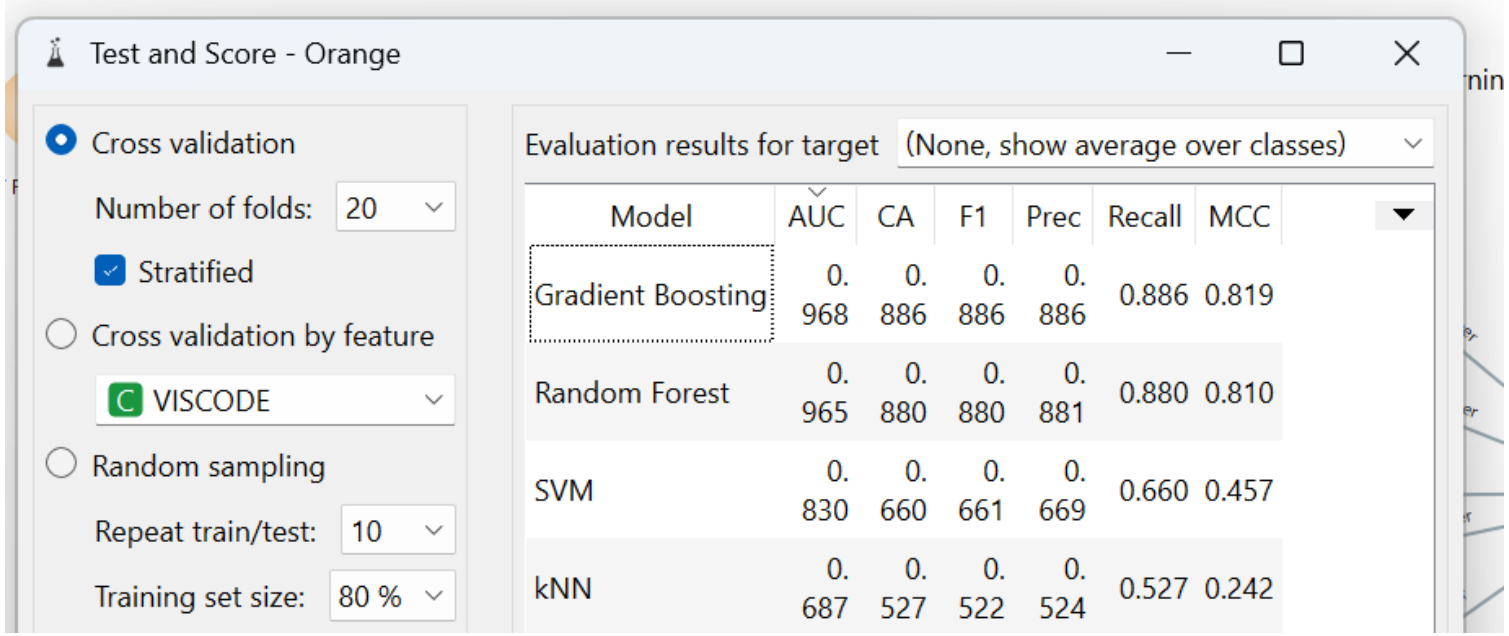




ADNI Workflow in Orange

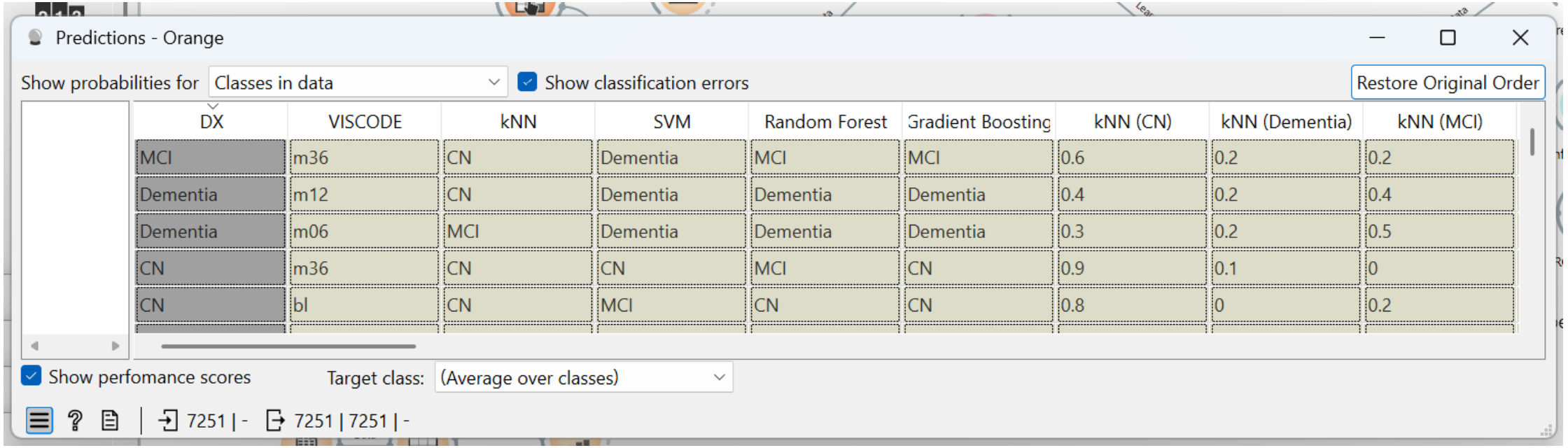
- 3 – Model Training:





ADNI Workflow in Orange

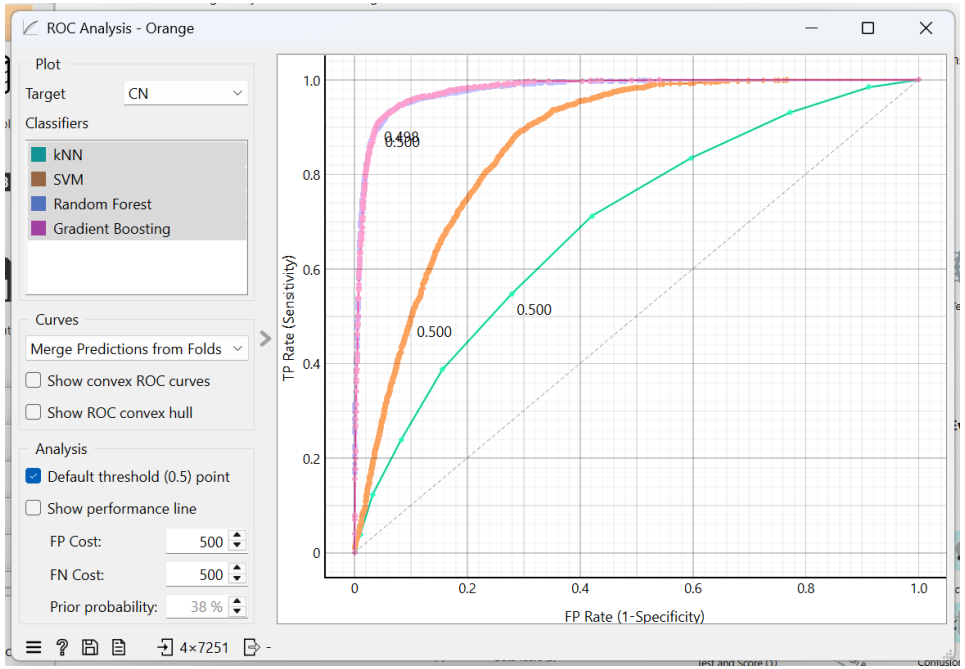
- 4 – Model Evaluation:



DX	VISCODE	kNN	SVM	Random Forest	Gradient Boosting	kNN (CN)	kNN (Dementia)	kNN (MCI)
MCI	m36	CN	Dementia	MCI	MCI	0.6	0.2	0.2
Dementia	m12	CN	Dementia	Dementia	Dementia	0.4	0.2	0.4
Dementia	m06	MCI	Dementia	Dementia	Dementia	0.3	0.2	0.5
CN	m36	CN	CN	MCI	CN	0.9	0.1	0
CN	bl	CN	MCI	CN	CN	0.8	0	0.2

Model Performance and Interpretation

- Visualize model performance



		Predicted			
		CN	**Dementia**	**MCI**	**∑**
Actual	**CN**	2509	0	243	**2752**
	Dementia	2	1131	219	**1352**
	MCI	195	171	2781	**3147**
	∑	**2706**	**1302**	**3243**	**7251**

Try it Yourself!

- Accessible AI without programming
- Fast data exploration and model prototyping
- Intuitive understanding via visual workflows

https://orangedatamining.com/

- Try it out with our sample workflows and synthetic datasets!