



UNIVERSITY OF
ARKANSAS

Inspire Finance Students of Machine Learning

-The visual way

Cinder (Xinde) Zhang
University of Arkansas

[FMA 2022 Atlanta](#)



The Demand

- CNBC 2019
 - ***“Finance jobs requiring A.I. skills increased 60% last year”***
- CFI list Machine Learning as one of the top 6 Skills in Finance



The Challenge

- Finance Students
 - Little or no coding experience
 - Uncomfortable with data and modeling
 - And more...
- Educators
 - Lack of effective and intuitive tool for teaching the subject





UNIVERSITY OF
ARKANSAS

Visual Approach

- Process
 - Define a meaningful business research question
 - Research data availability
 - Design the process
 - Data collection/processing
 - Feature generation
 - Build machine learning model(s)
 - Model testing
 - Refine the model and/or question
- The Tool
 - Dataiku
 - Assemble the process visually
 - With the steps as visual blocks



Example

- The Question: Can we identify factors which may predict Bitcoin (BTC) daily return?
 - First guess
 - Active address of BTC?
 - Data and resources
 - <https://github.com/CinderZhang/FinML/blob/main/ML%20Resource.ipynb>
 - [Metric Catalog - Glassnode Studio - On-Chain Market Intelligence](#)

[illegible]



UNIVERSITY OF
ARKANSAS

Building Blocks


1: Datasets

Datasets


Search

All

3 datasets



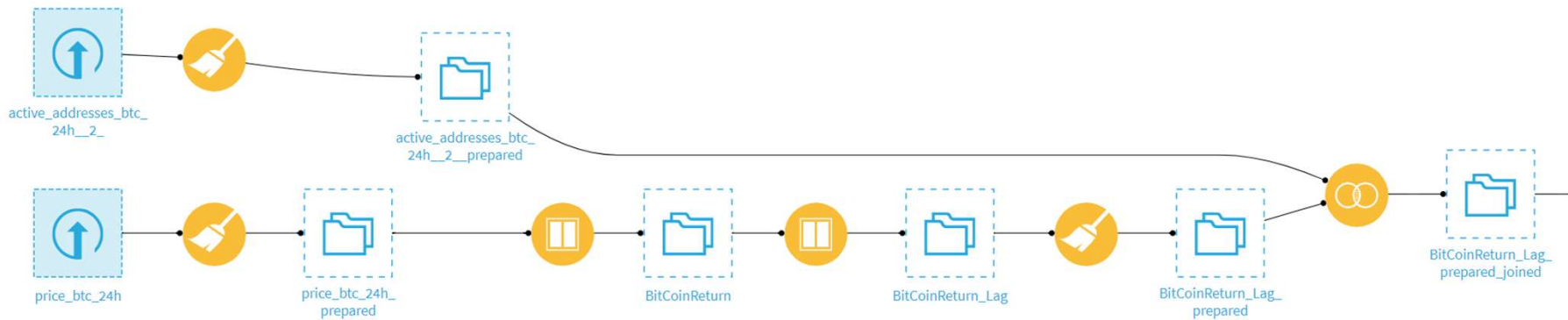
price_btc_24h



active_addresses_btc_
24h__2_

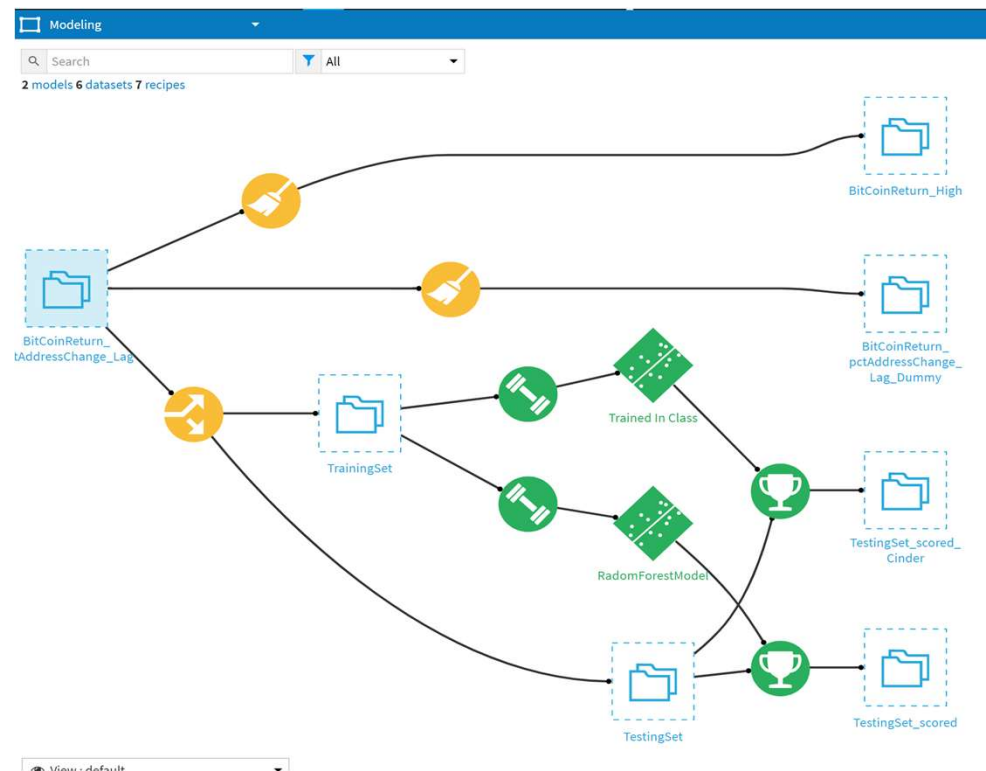
Building Block 2

2: Data Processing



Building Block 3

3: Modeling and Testing



Close Look of the Model(s)

XindeZhang_Blockchain_Testing Visual Analyses

Quick modeling of Return on TrainingSet / Models / Random forest

Summary

What if?

EXPLAINABILITY

Decision trees

Variable importance

Partial dependence

Subpopulation analysis

Individual explanations

PERFORMANCE

Scatter plot

Error distribution

Metrics and assertions

MODEL INFORMATION

Data preparation

Features

Algorithm

Hyperparameter optimization

Training information

Features

Input features

Search... 1-14 of 14

Price_lag	Rejected	# Numeric	
Return	Target	# Numeric	
NewAddress_lag	Rejected	# Numeric	
ActiveAddress_lag	Input	# Numeric	Avg-std rescaling
Price	Rejected	# Numeric	
t_month	Input	# Numeric	Avg-std rescaling
t_year	Input	# Numeric	Avg-std rescaling
NewAddress	Rejected	# Numeric	
t_day	Input	# Numeric	Avg-std rescaling
Date	Rejected	# Numeric	
NewAddress_pct_lag	Input	# Numeric	Avg-std rescaling
ActiveAddress	Rejected	# Numeric	
NewAddress_pct	Rejected	# Numeric	
Price_lag_diff	Rejected	# Numeric	



Future Work for Students

- More data
- More possible factors
- Refine the Question



UNIVERSITY OF
ARKANSAS

Ah!





UNIVERSITY OF
ARKANSAS

Thank you

