1. Description of Data Set

The topic of this data set is to predict the success of an investment in a cryptocurrency token and coins (called tokens in the following). As the prediction of stock returns is a classical application of predictive analytics, predicting token performance is an emerging field of application. A precise prediction of successful tokens is of great value for investors, corporate representatives, and researchers.

This data set contains of various data points representing each token: Company information, links to social media and development channels, market information, community feedback, and information about software development, the Initial Coin Offering (ICO), and other publicly available information.

The value of a high-quality prediction depends on the perspective of stakeholder. While investors in cryptocurrencies seek to maximize their returns, corporate representatives evaluate whether an ICO might be an alternative to an Initial Public Offering (IPO), and researchers are interested to better understand an emerging phenomenon.

2. Classification Task

This data set contains token data including historic price points from January to September 2019. Within the scope of the data set, a successful investment is defined as a collection of different criteria. Most important a token must not be a scam and positive returns are desired. Your task is to predict whether an investment in a token will be successful or not.

It is permitted to use external data within the following limited scope. It may be interesting to integrate the current and historic bitcoin price into your analysis. It may also be worth having a look at the provided links in the links_* section to enrich the token data set.

3. Feature Overview: Token Data Set

Feature	Example	Description
OBS_ID	235691	Observation-ID, ascending
Id	02-token	individual id for each token
Symbol	o2t	A symbol (3 – 6 letters) the token is traded under
Name	O2 Token	The name of the token
categories_0	„Business Services"…	An assigned category
contract_address	0xb1bafca3737…	Contract address of the token on the Ethereum blockchain
transaction_count	37685	Total number of transactions, where this token was involved

holder_count	452	Total number of current token holders
timestamp	Dec-10-2018 03:39:32 AM	The date of the last transaction
country_origin	"US", "SI", "FR", …	Two letter code for the country of origin
links_*		
*homepage	http://o2token.io/	
*blockchain_site	etherscan.io/token/…	URL to additional information about this token
*twitter_screen_name	0xcert	
*facebook_username	0xcert	
*subreddit_url	www.reddit.com/r/0xcert	
*repos_url_github	github.com/0xcert	
market_data_*		
*current_price_usd	0.00083971	
*ath_usd	0.03418169	All time highest price (All Time High, ATH)
*ath_date_usd	2018-10-10T17:27:38.034Z	Date of the ATH
*high_24h_usd	0.00168653	Highest price in the last 24 hours
*low_24h_usd	0.00167661	Lowest price in the last 24 hours
*total_supply	65658824.0	Total number of tokens
*circulating_supply	8627701.3340348	Tokens that are available for trading
*first_time_trading	Sep-15-2018 02:00:00 AM	
community_data_*		
*facebook_likes	25	
*twitter_followers	345	
*reddit_...		Various reddit metrics
*telegram_channel_user_count	11236	
developer_data_*		
*forks	3	
*stars	654	
*subscribers	45454	
*total_issues	45	
*closed_issues	34	
*pull_requests_merged	12	

*pull_request_contributors	3	
*commit_count_4_weeks	6	
ico_data_*		
*amount_for_sale	250000000.0	Tokens available for sale in ICO
*country_origin	SI	
*hardcap_amount	20000.0	
*hardcap_currency	ETH	
*kyc_required	True	Is "Know Your Customer" required?
*pre_sale_available	True	Was a pre-sale available
*softcap_amount	5000.0	
*softcap_currency	ETH	
*total_raised	9450000.0	
*total_raised_currency	USD	
public_interest_stats_*		
*alexa_rank	10926950	www.alexa.com/topsites Alexa rank of token homepage
*bing_matches	23900	Number of Bing search results for token name
KW1 – KW39	0.003453	A historic price point for each calendar week 2019
Success	**0 := unsuccessful** **1 := successful**	**Dependent variable**

4. Feature Overview: Bitcoin Price Data Set

Feature	Example	Description
date_in_ms	1367107200000	Is the number of seconds that have elapsed since the Unix epoch, that is the time 00:00:00 UTC on 1 January 1970, minus leap seconds
price_usd	135.3	The bitcoin price in US Dollar