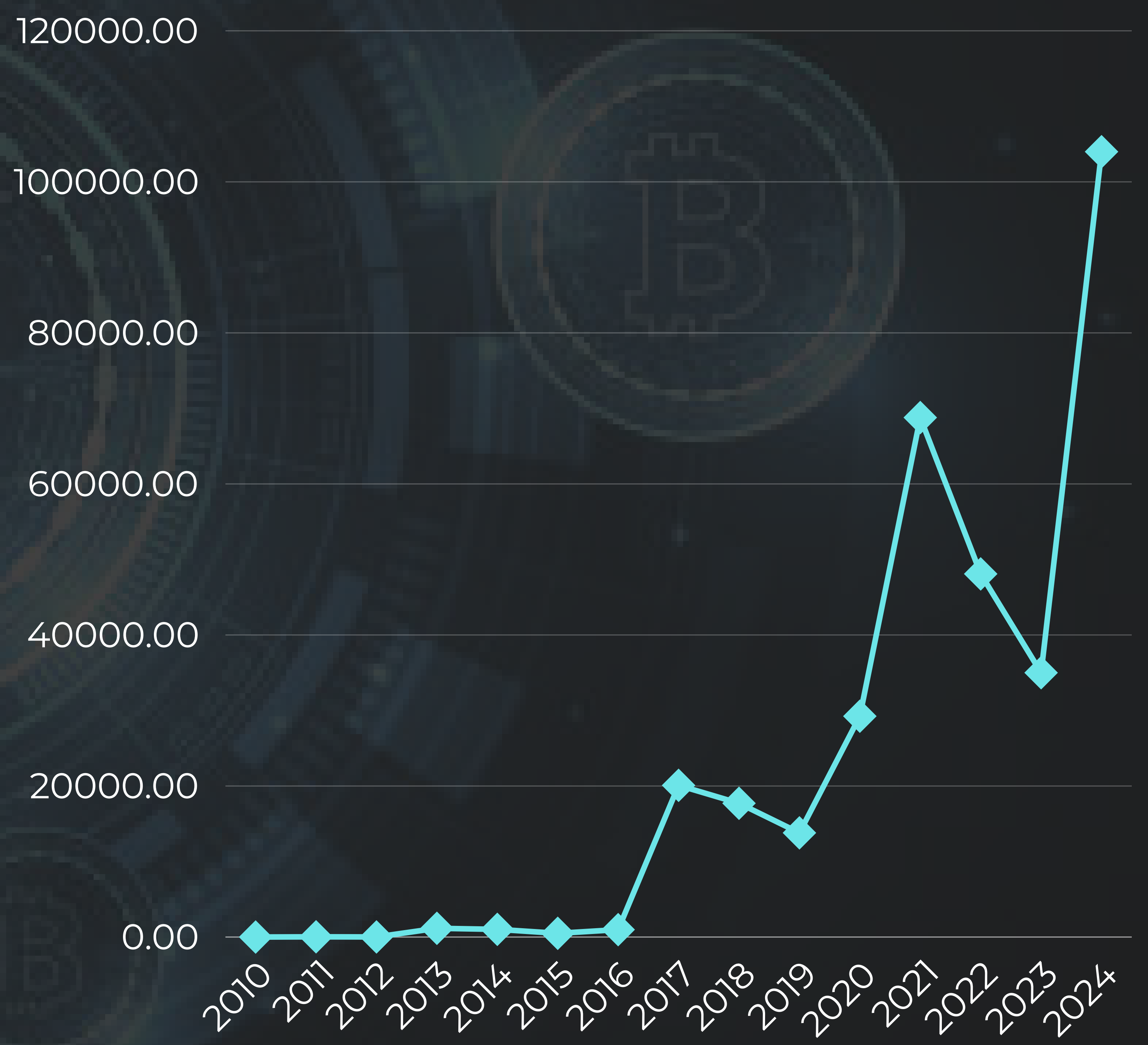
CALIFORNIA STATE UNIVERSITY
LONG BEACH
College of Business

Analyzing Factors Influencing Cryptocurrency Investment

RISE OF CRYPTOCURRENCY

- Cryptocurrencies emerged in 2009 with Bitcoin, growing from obscure digital tokens to a multi-trillion-dollar market.
- Extreme price volatility, regulatory uncertainty, and frequent security breaches expose investors to severe risks.
- Many coins have no intrinsic value, and their energy-intensive operations raise environmental concerns.



PROBLEM STATEMENT

PROBLEM

Despite the growing popularity of cryptocurrency as an investment option, many individuals invest without a comprehensive understanding of its **RISKS**, **TECHNOLOGY**, or **MARKET DYNAMICS**.

IMPACT

This lack of knowledge can lead to **UNINFORMED DECISIONS**, **FINANCIAL LOSSES**, and increased vulnerability to market volatility.

SCOPE

By analyzing survey data, this study aims to identify the factors that influence individuals to invest in cryptocurrency and provide insights to improve **FINANCIAL LITERACY** and decision-making.

DATASET

American Trends Panel Wave 111
from PEW research centre

- Shopping habits
- Payment methods
- Social media usage
- Online dating
- Usage of cryptocurrencies
- Knowledge of NFTs
- Sports betting

DATA PRE-PROCESSING

1. RAW DATA
2. HANDLING MISSING DATA
3. RECODE VARIABLE
4. COMBINE AND SPLIT VARIABLES
5. CHECK FOR OUTLIERS
6. DATA VALIDATION
7. CLASSIFY TARGET AND FEATURES

KEY VARIABLES

Target Variable

Crypto_Invest

A binary variable with values indicating whether the respondent has invested in cryptocurrencies

Feature Variables

- Online_Shopping_Phone
- Prefer_Shopping_Online
- Influenced_Purchace
- Married
- Used_Dating_Site
- Tinder
- Cash_App
- Pay_Pal
- Payment_Account_Hacked
- Online_Shopping_Freq
- Gambling_Know
- NFT_Invest
- Online_Betting
- Rural
- Metropolitan
- Male
- Age
- Education
- NFT_Know
- Crypto_Know

DATASET

DATASET A
Strict_Cleaned_Data

All rows with missing values or invalid values removed

DATASET B
Threshold_Cleaned_Data

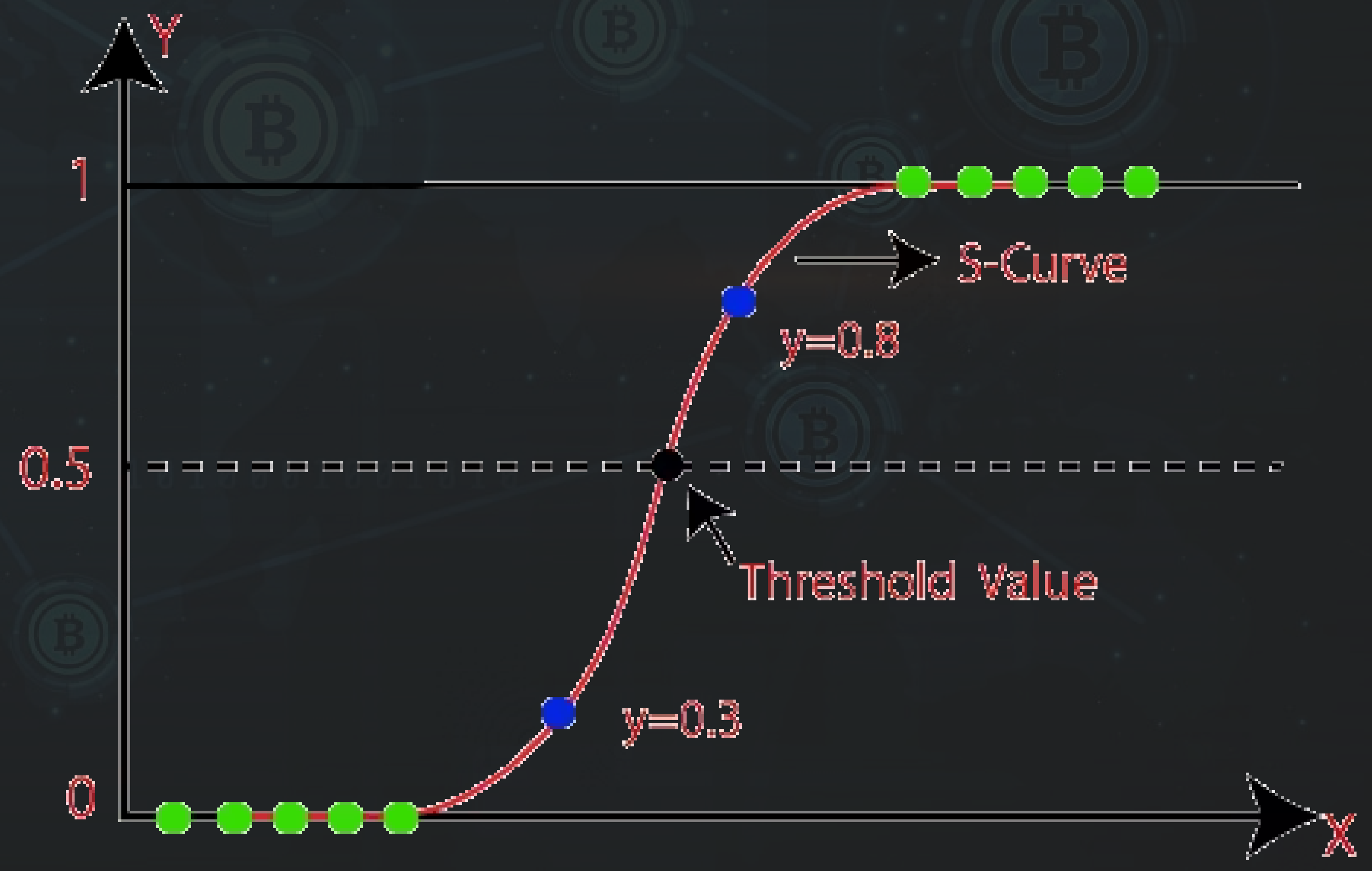
Rows if it has >30% of missing values removed

MODEL USED

Logistic Regression model

- This model shows the relationship between one or more independent variables and a dependent variable (outcome) using a log-odds transformation.

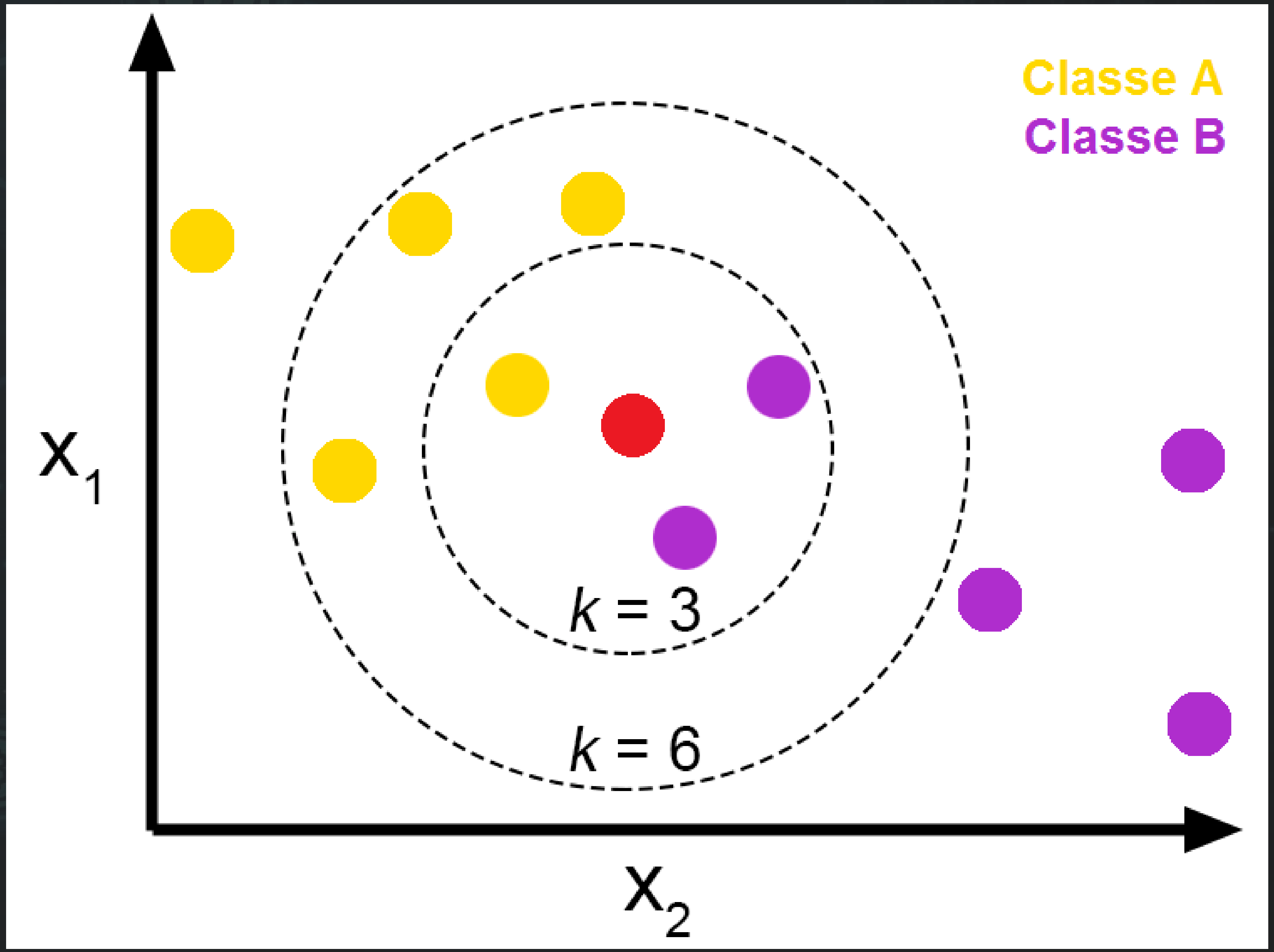
- It identifies how the likelihood of the dependent variable changes with variations in the independent variables.



K-Nearest Neighbor (KNN)

- The dependent variable can be categorical, and KNN assigns the most common class among the K-nearest neighbors.

- The dependent variable is numerical, and KNN predicts its value by averaging the values of the K-nearest neighbors.



MODEL TRAINING AND VALIDATION

MODELS USED:

- Logistic Regression: A simple and interpretable model ideal for binary classification.
- K-Nearest Neighbors (KNN): A distance-based model that considers the closest data points.

TRAINING PROCESS:

- Dataset split into training and validation sets to ensure balanced evaluation.
- Focused on minimizing overfitting for generalization.

INPUT DATA → DATA CLEANING → FEATURE ENGINEERING → MODEL TRAINING → VALIDATION

RESULTS

SUBSET 1:
- **Logistic Regression - 77.2%**
- **KNN - 75.2%**

SUBSET 2:
- **Logistic Regression - 84.8%**
- **KNN - 83.6%**

CALIFORNIA STATE UNIVERSITY
LONG BEACH
College of Business

SIGNIFICANCE of FEATURES

ALPHA LEVEL = 0.05

<0.05

1. Cash_App
2. Crypto_Know
3. NFT_Invest
4. Online_Betting
5. Male
6. Tinder
7. Prefer_Shopping_Online
8. Rural
9. Metropolitan
10. OkCupid
11. Gambling_Know

<0.1

1. Me_too_Support
2. Age
3. Match
4. Grindr
5. Liberal
6. Age_30_49
7. Pay_Pal
8. Age_18_29

<0.5

1. Hinge
2. Venmo
3. Zelle
4. Influenced_Purchase
5. Online_Shopping_Phone
6. Middle_Income
7. Bumble
8. Inperson_Betting
9. Influenced_by_Influence
10. Constant
11. eharmony
12. Sub_Urban
13. Single

>0.5

1. Education
2. Family_Income
3. Married
4. Online_Shopping
5. Online_Shopping_Freq
6. Payment_Account_Hack
7. Payment_Site_Scam
8. Upper_Income
9. Urban
10. US_Citizen
11. Use_Social_Media

INSIGHTS

MEN ARE
93%
MORE LIKELY TO INVEST IN CRYPTOCURRENCY.

CRYPTO_INVEST

USERS OF CASH APP ARE
92.3%
MORE LIKELY TO INVEST IN CRYPTOCURRENCY

PEOPLE IN MID-INCOME TIER INVEST MORE IN CRYPTO THAN PEOPLE IN UPPER-INCOME TIER

MARRIED people are **17%** more likely to invest in cryptocurrency than UNMARRIED people

PEOPLE WHO USE ONLINE PAYMENT APPS ARE MORE LIKELY TO INVEST IN CRYPTO

HAVING KNOWLEDGE OF CRYPTOCURRENCY INCREASES THE ODDS OF INVESTMENT BY
319.7%

USE CASES

USE CASE 1

Policy Development and Consumer Protection

USE CASE 2

Risk Assessment for Financial Institutions

CONCLUSION

This project identified key factors influencing cryptocurrency adoption, such as knowledge of cryptocurrency, online betting, and NFT investments. Men and suburban residents were more likely to invest, while behavioral and demographic patterns played a significant role. To improve, the study can incorporate larger datasets, address prediction bias, and include real-time financial trends to enhance model accuracy and applicability. These findings can guide financial platforms and policymakers in effectively targeting potential investors.