

Presented by Elie Niringiyimana

Consumer survey

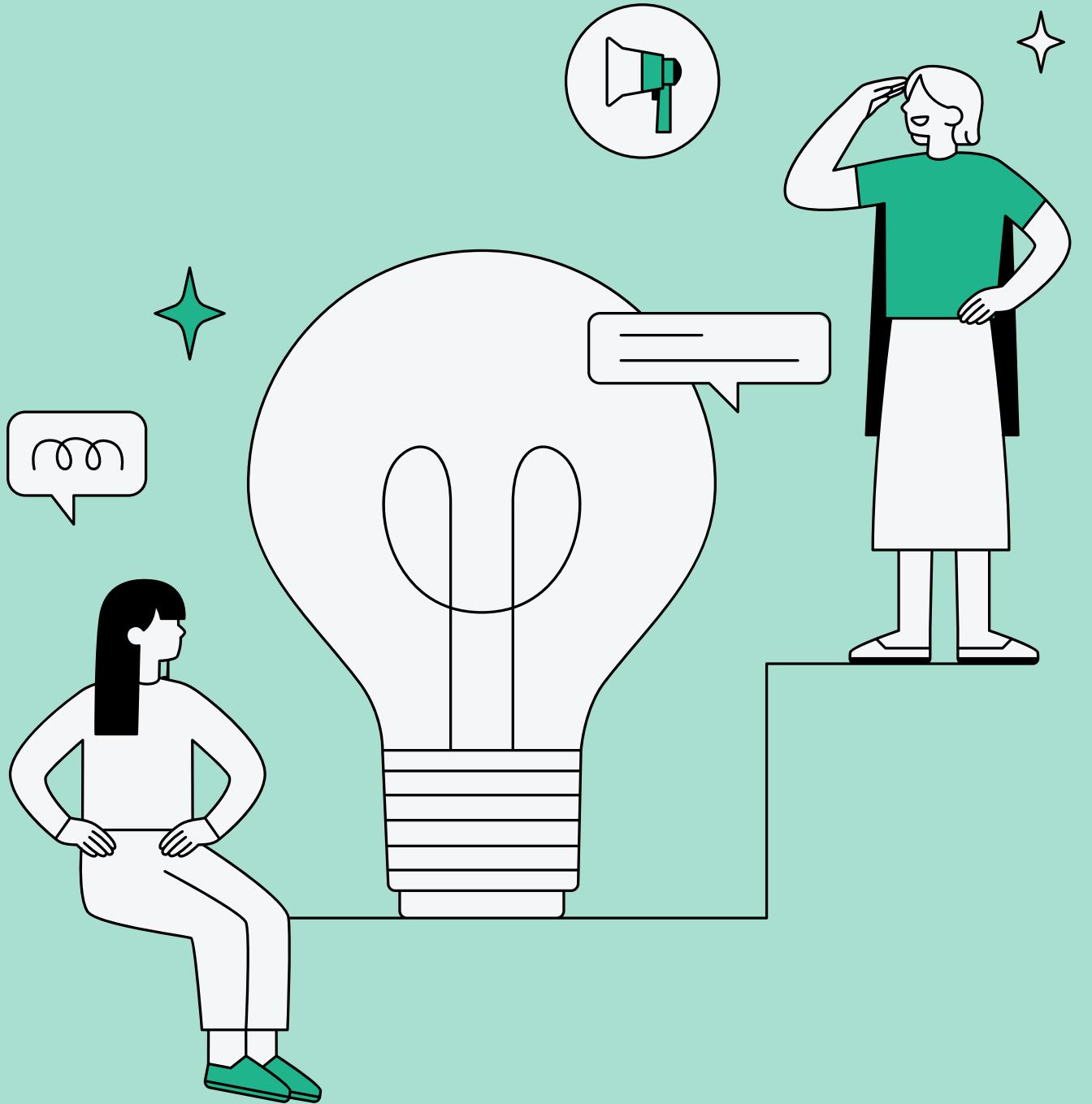
Automating handwritten number  
recognition

1



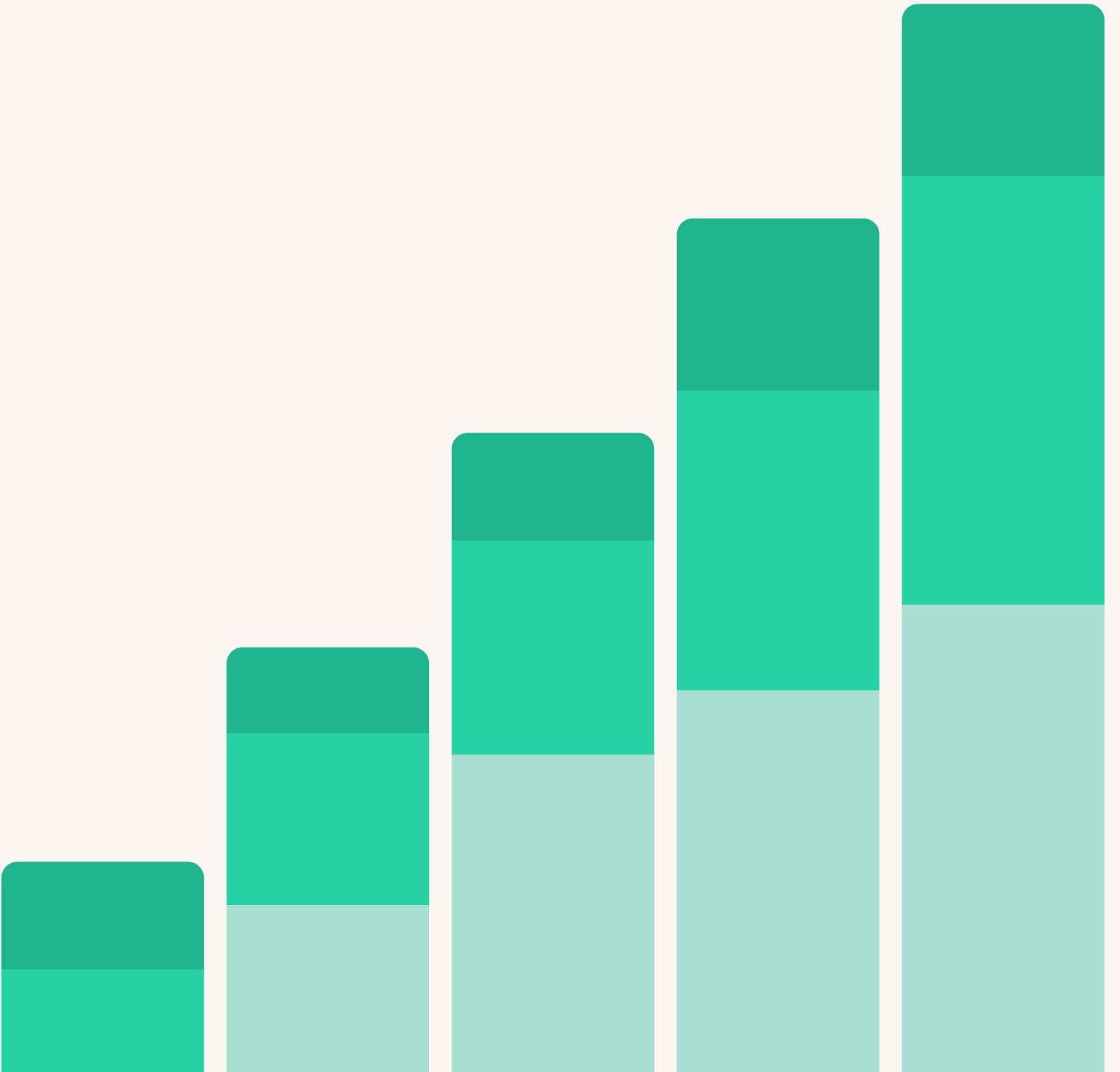
## Background on the Consumer Survey Process

- Traditional consumer surveys often involve handwritten responses, which are manually entered into systems.
- This manual entry is time-consuming, prone to errors, and inefficient.
- The goal is to automate the recognition of handwritten numbers to streamline this process.



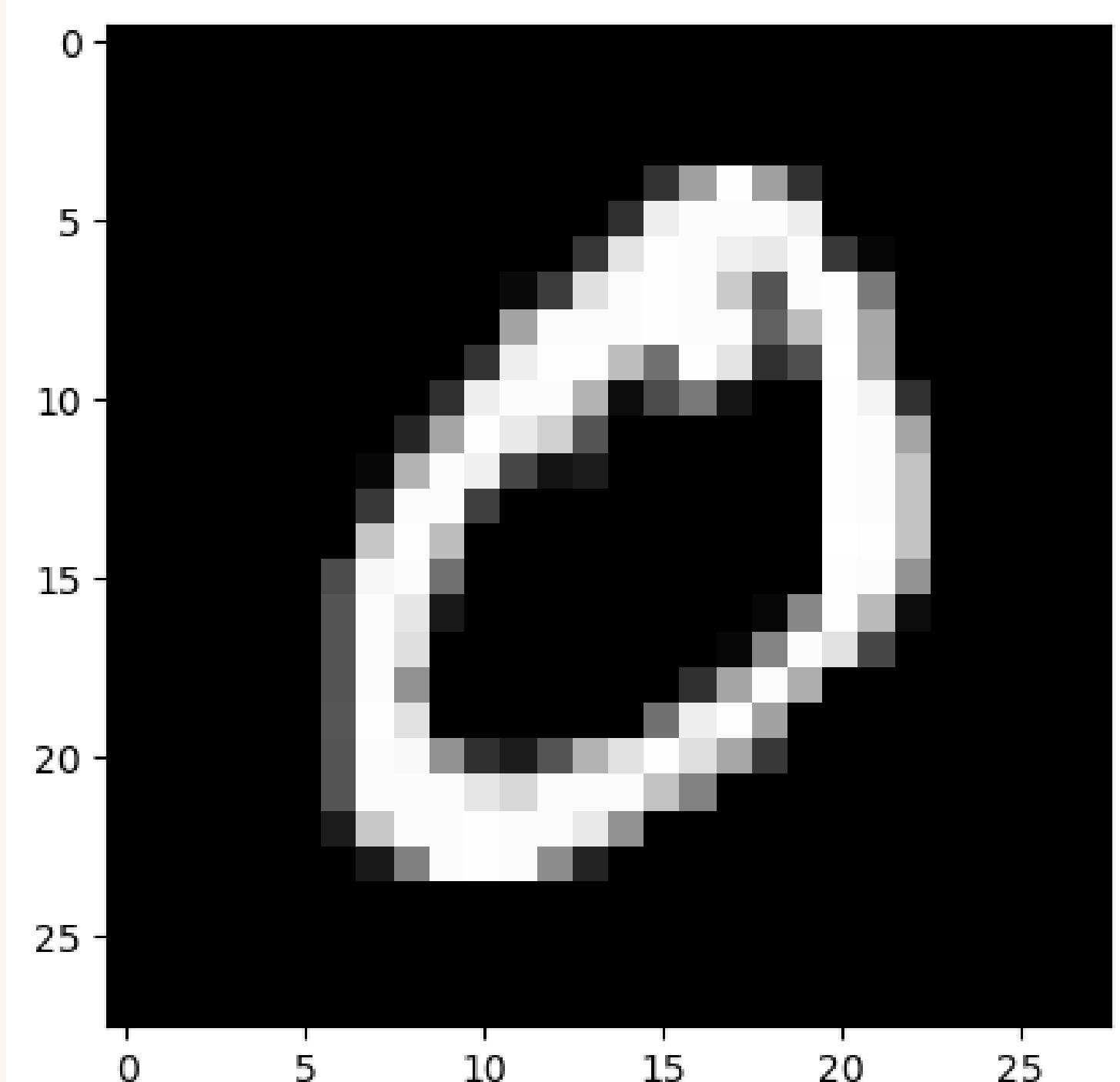
## Proposed Solution

- Implement a machine learning model using the MNIST dataset.
- The model will be trained to recognize handwritten numbers.
- This solution will automate the extraction of handwritten numeric data from surveys.



## Data Understanding

- The MNIST dataset consists of 60,000 training images and 10,000 testing images of handwritten digits (0-9).
- Each image is 28x28 pixels, with grayscale intensity values.
- Data type: int64
- No missing values
- No duplicates



## Data Preparation

- Reshaping and normalizing the data for model input.
- Addressing class imbalances by implementing under sampling.

## Modeling

- The K-Nearest Neighbors (KNN) algorithm with 5 neighbors
- The model was trained using MNIST train dataset.
- The model was tested using MNIST test dataset.

## Model Evaluation

The accuracy is: 96.9%



## Use Case Implementation

- The trained model can be integrated into a system to automatically process handwritten numeric data from surveys.
- Improve data processing speed enables quicker decision-making.
- Automate data entry from handwritten surveys saves time and reduces errors.

# Summary

- Used numpy and pandas to model hand written number recognition.
- Reshaped and normalize data for model input
- Addressed class imbalance by implementing under sampling technique.
- Trained the model with KNN algorithm with 5 neighbors.
- The model resulted 96% of accuracy.



Presented by Elie Niringiyimana

Thank you very much!

10

