Empowering Trustworthy Online Shopping

RealView App

E-commerce in 2022-2023

12%

Expansion

2023 E-commerce sales grew 12% from 2022

15.4%

Penetration

E-commerce sales accounted for 15.4% of all sales

2.6B.

Resilience

Estimated 2.6 billion digital buyers. Only continuing grow during the pandemic

Why reviews are so important?

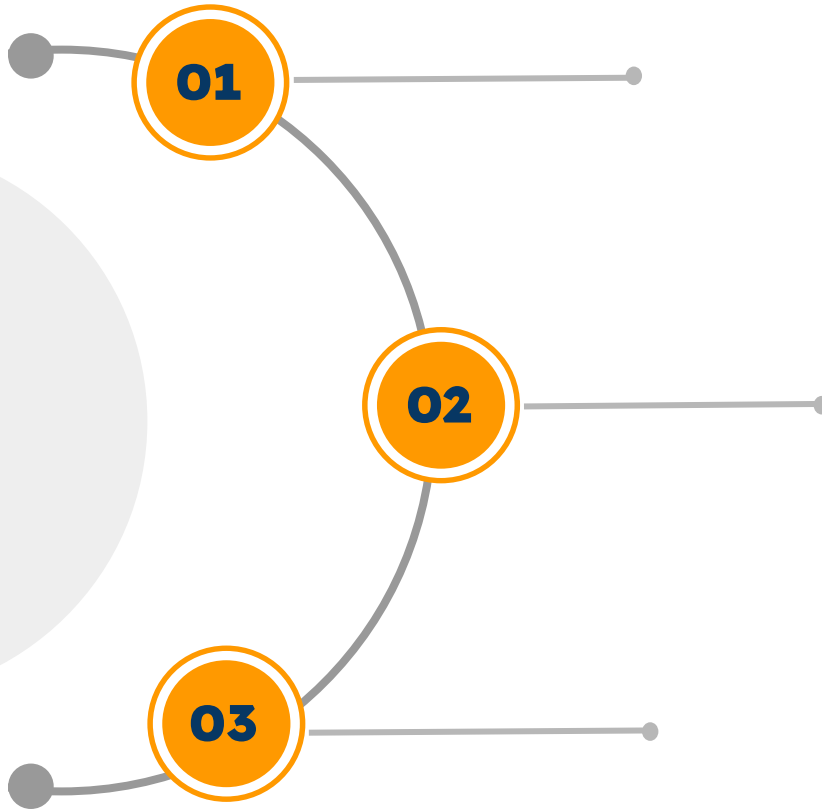**270%** | Products with reviews have a 270% higher conversion rate

89% | 89% of customers read reviews before buying a product

75% | 75% people trust online reviews as much as personal recommendations

Business problem

- 40% of online reviews are fake
- 90% of these fakes are AI-generated, leading to misleading consumer decisions

Stakeholders

- Consumers seeking reliable information before making online purchases
- E-commerce platforms and retailers looking to maintain trust and reputation with their customers

Solution

- We created the 'RealView' model
- It will empower users with credible reviews for better shopping decisions

01

02

03

Data Explanation

- **The data source from OSF.IO comprises reviews related to online shopping**

20260 fake reviews

Computer-generated reviews are artificially created feedback produced by automated processes

20260 real reviews

Authentic reviews are genuine feedback provided by real individuals, reflecting their actual experiences with a product or service

After exploring the data, we found a satisfyingly consistent Label Distribution

Distribution of Ratings for Different Labels

CG

OR

Ratings legend:
- 5
- 4
- 3
- 2
- 1

The sole distinction we identified was in the average word length (word count), with **73.64** for Organic (OR) and **61.29** for Computer-Generated (CG)

- We observe balanced ratings for both Computer-Generated (CG) and Organic (OR) labels.

Ensemble Methods

Best Cross-Validation
Score: 0.727

Best Cross-Validation
Score: 0.726

Best Cross-Validation
Score: 0.779

Logistic Regression Model	Random Forest Model	AdaBoost Classifier	Gradient Boosting Classifier	XGB Classifier

Best Cross-Validation
Score: 0.762

Best Cross-Validation
Score: 0.738
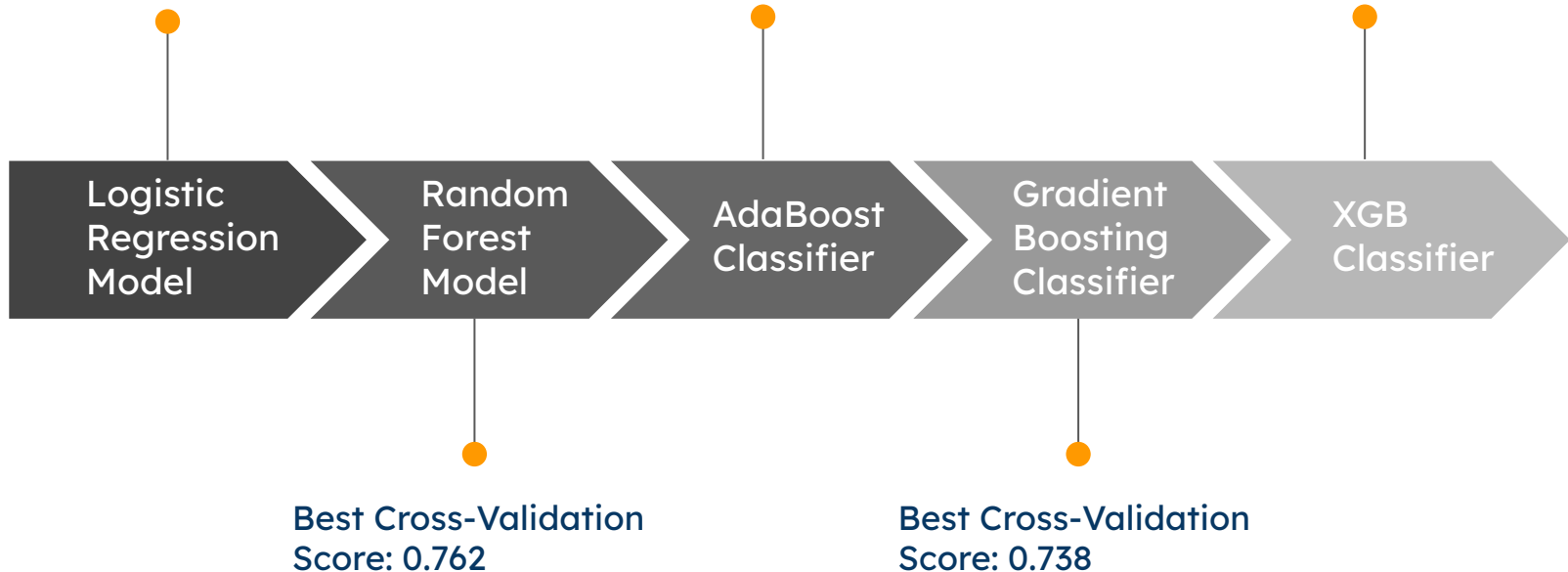
Performance of the Best Neural Network Model

Confusion Matrix for Deep NN Model

	Predicted 0	Predicted 1
True 0	5042	1023
True 1	1646	4420

Accuracy history
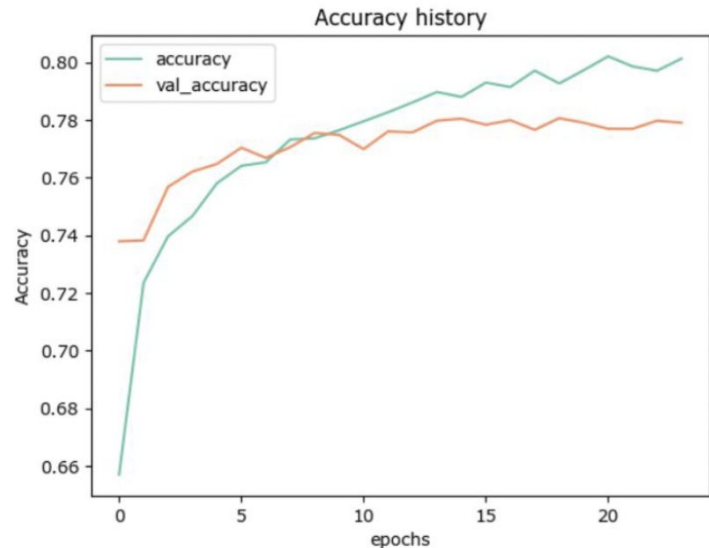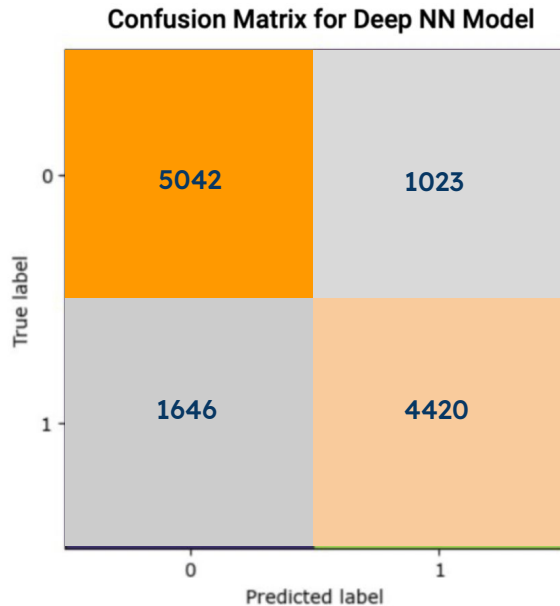
- The deep Neural Network achieved a test accuracy of 0.78, outperforming all previous models

Future Work

- Build **RealView** App

- Develop real-time review detection capabilities

- Create a Feedback Mechanism, that enables users to report suspicious review

- Extend the reach of **RealView** to international markets and multiple languages

Contacts

Eduardo Gonzalez
GitHub: https://github.com/0xEduardoG
LinkedIn: https://www.linkedin.com/in/eduardo-dgonzalez/
Email: eduardo0591@gmail.com

Ekaterina Serbina
GitHub: https://github.com/serbinaekaterinai
LinkedIn: https://www.linkedin.com/in/ekaterina-serbina/
Email: serbinaekaterina99@gmail.com