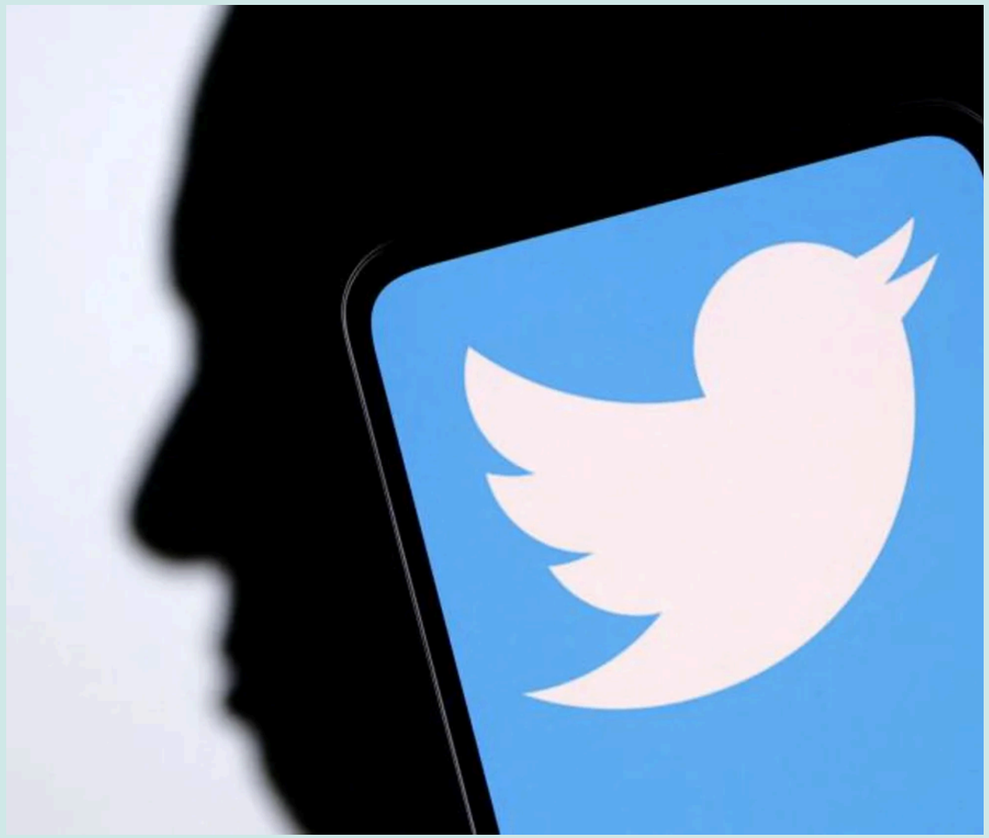
PSYCHOPATHY PREDICTION ON TWITTER

PRESENTED BY GROUP 14

Juan David Bejarano Cristancho 20232020056
Raúl Andrés Díaz Losada 20232020058
Juan David Avila 20232020154
David Sánchez Acero 20232020049



Can we understand the human mind through social media?

- 2,927 users – 80 countries – 3 million tweets
- 337 linguistic, behavioral, and social features
- Goal: Understand psychological traits through online behavior

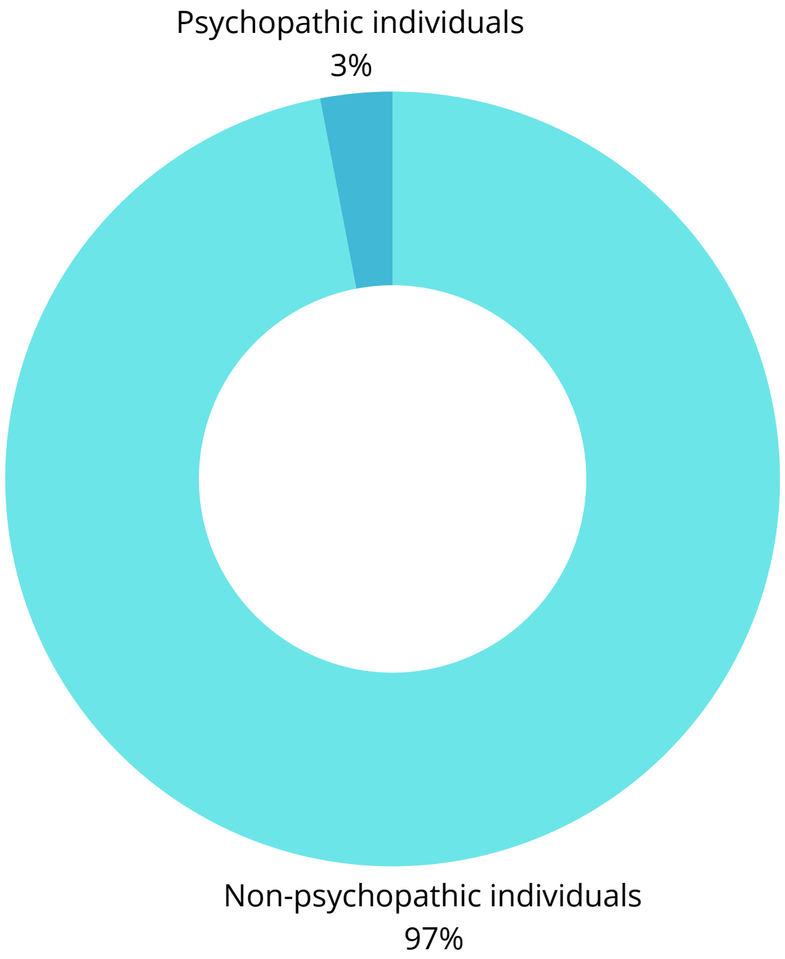
The 3% problem

Unbalanced data

very few high-psychopathy cases, this could produce something called “class imbalance”

Predictions

Hard to create fair or accurate predictions



Understanding the system

Twitter Data → Feature Engineering → Modeling → Evaluation

Interconnected subsystems

All parts of the system depend on each other — data collection, processing, modeling, and evaluation. A small issue in one stage can affect the rest, showing the need to view the system as a whole.

Information feedback loops

The system constantly reacts to its own output. In Kaggle, leaderboard results influence participants' behavior, which then changes future results, creating a self-adjusting cycle.

Sensitivity to small changes

Tiny variations in data or parameters can cause big shifts in results. This shows the system's chaotic nature — stable in structure, but unpredictable in behavior.

Chaotic and Emergent Behavior

- High sensitivity to initial conditions
- Feedback loops: Rankings ↔ Participants
- Emergent patterns not planned in advance

Predictable rules,
unpredictable outcomes

From Analysis to Design

A More Stable System

- Data Ingestion – collect clean, anonymous input
- Feature Processing – convert text into measurable variables
- Modeling – learn and predict using structured data
- Evaluation – measure performance fairly
- Deployment & Feedback – continuous monitoring

Engineering Principles:

- Modularity
- Scalability
- Reliability
- Maintainability

Handling Chaos and Sensitivity

The goal in this stage was not to eliminate uncertainty but to manage it. We proposed several strategies that would help stabilize a complex and sensitive system, such as keeping data standardized, combining multiple models to balance errors, and continuously monitoring results to detect sudden changes.

Beyond technical stability, we also considered ethical and privacy aspects, because a system that deals with human data must be responsible and transparent. The idea was to design a framework that remains adaptable, resilient, and trustworthy even when exposed to unpredictable behaviors.

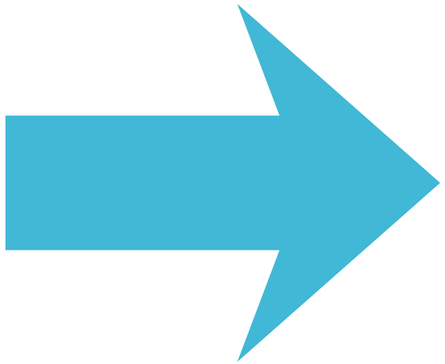
- Data standardization
- Ensemble methods (combine models)
- Continuous performance monitoring
- Ethical and privacy awareness

“Not to eliminate uncertainty — but to manage it.”

Connecting Both Workshops

Workshop 1:

- Understood the system's complexity
- Identified chaos and sensitivity
- Observed emergent behaviors



Workshop 2:

- Proposed a robust architecture
- Focused on modularity and adaptability
- Addressed uncertainty and ethics

Challenges and Next Steps

Challenges:

- Data imbalance (3%)
- No real implementation yet
- Ethical & bias risks
- Human unpredictability

Next Steps:

- Build prototype version
- Integrate real-time data
- Explore federated learning
- Evaluate social impact

Thank you very much!