

07/07/2025

Dimitrov Yanik

# Analisi Predittiva dei Crimini a Chicago



...

# Problema

- Descrizione: «Vogliamo stimare la probabilità che si verifichi un furto ("THEFT") in ogni quartiere della città, usando dati storici sui crimini.»
- Utilità: «Può supportare le forze dell'ordine e le decisioni sulla sicurezza urbana.»
- Origine: «Il dataset originale contiene milioni di crimini avvenuti a Chicago.»

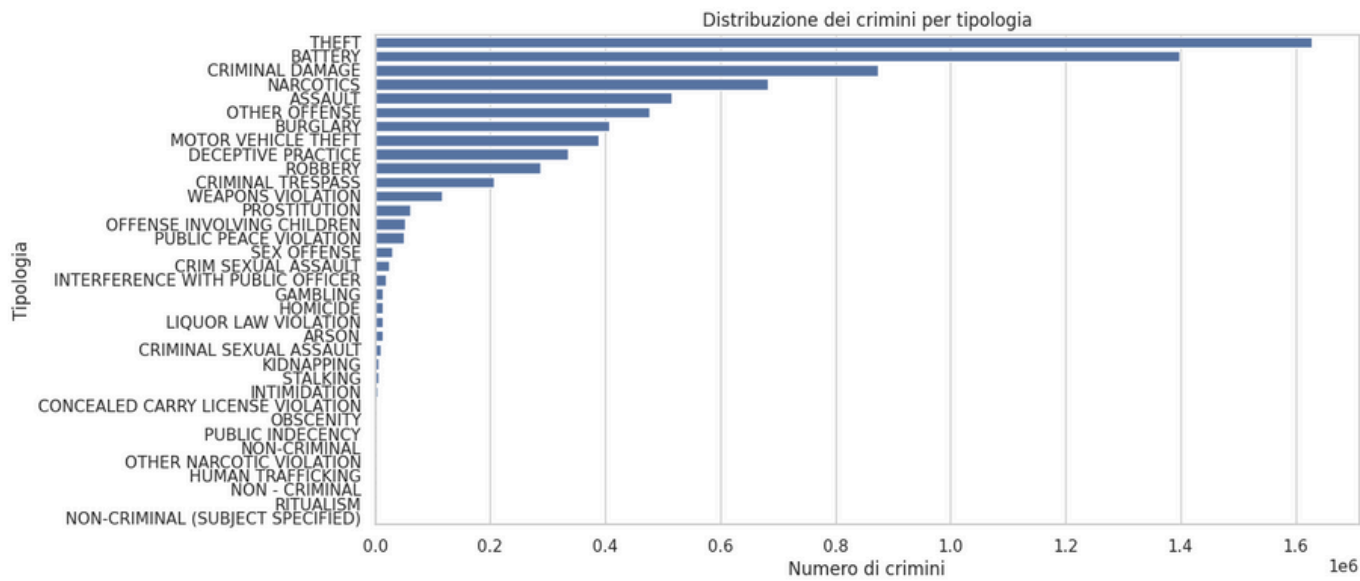
01

# 02

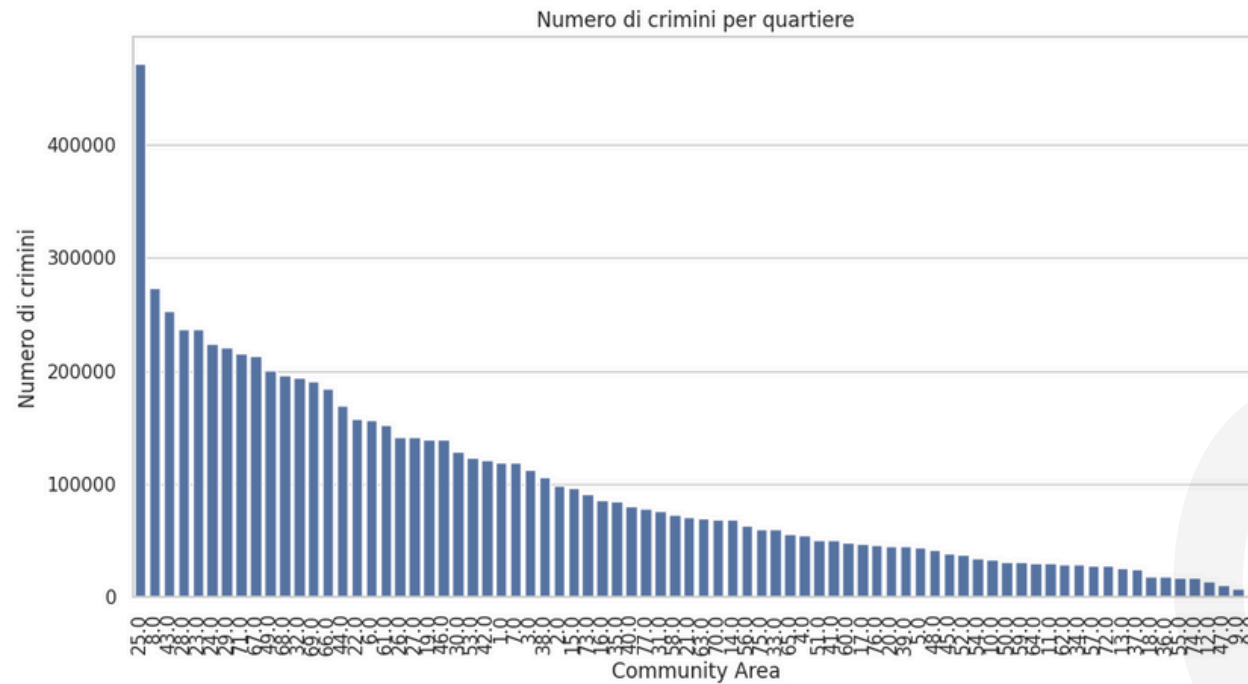
# Dataset

- Fonte: dataset pubblico del portale di Chicago
- Periodo: anni analizzati 2001 - Attuale
- Numero iniziale di record: oltre 7 milioni
- Variabili principali: Date, Primary Type, Community Area, Location Description, Latitude/Longitude

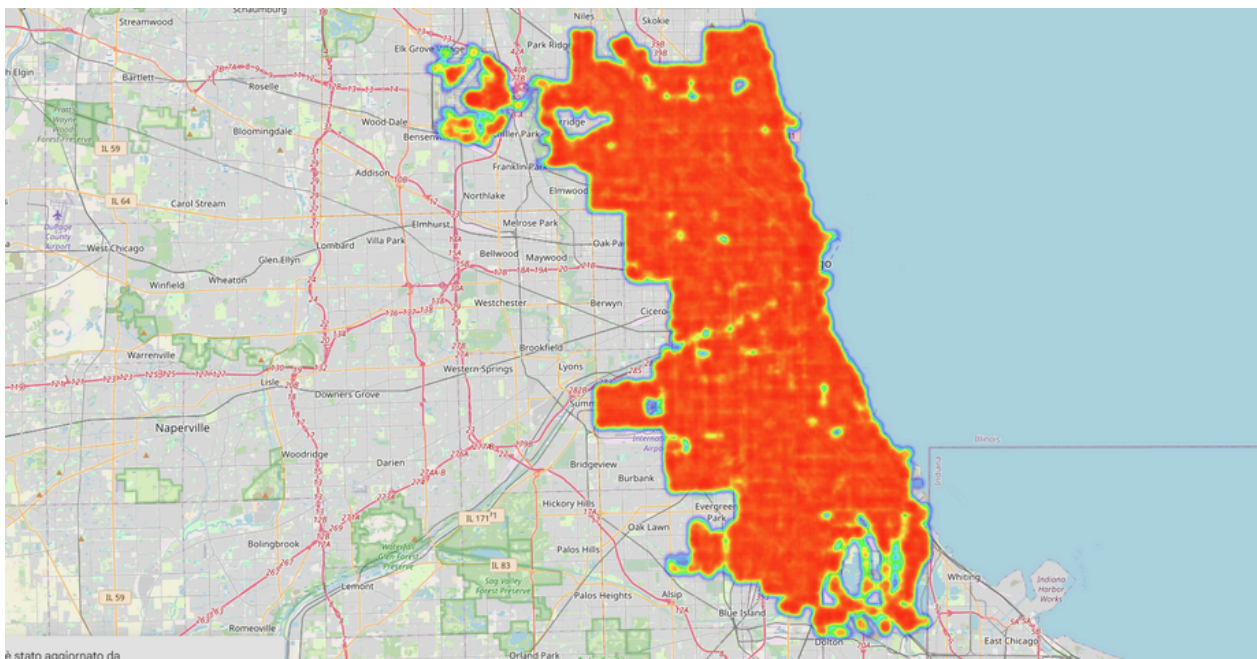
## Distribuzione dei crimini



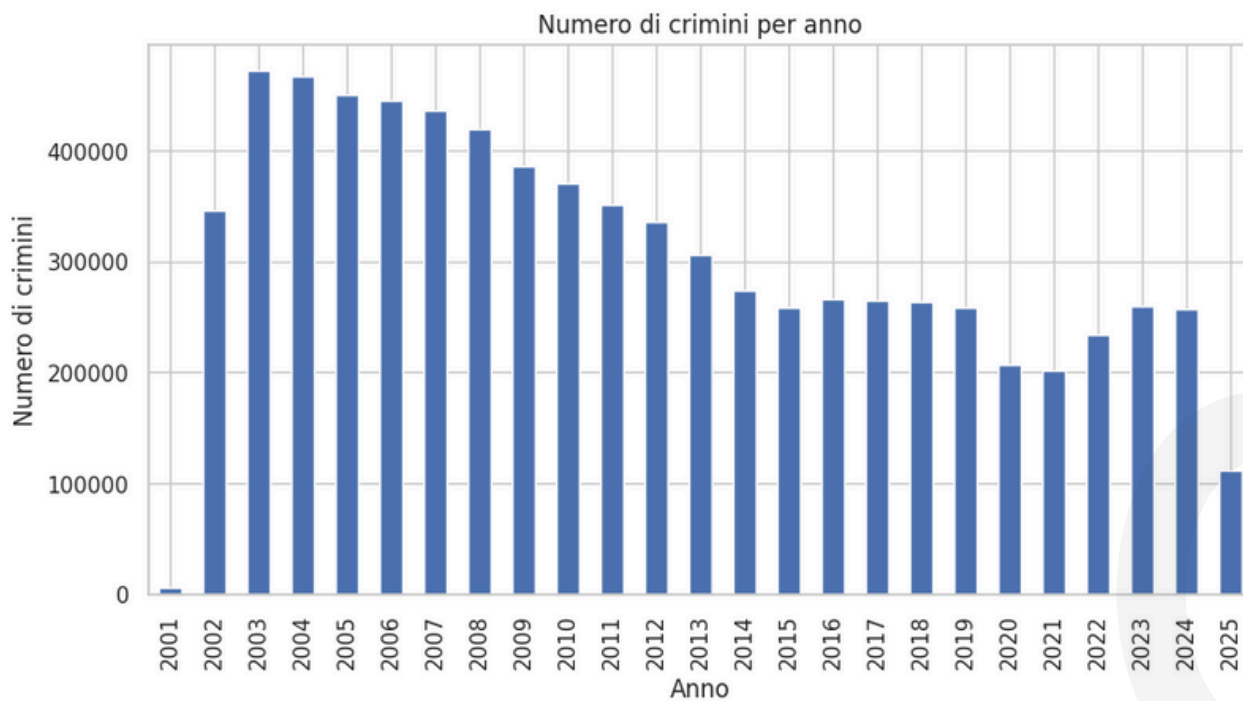
THEFT più frequente (~20%).



Community Area



Concentrazione nelle zone centrali.



Picchi in determinati anni.

Trend temporali  
e mappa

...

- Obiettivo: prevedere probabilità che un crimine sia un furto (THEFT).
- Algoritmo scelto: Random Forest Classifier.
- Perché? Robusto, gestisce dati numerici e categoriali, riduce rischio overfitting.

# Modello predittivo

- Feature usate:
- Community Area
- Location Description
- Year, Month, Hour

# Risultati

Analisi risultati

```
Accuracy: 0.7955
ROC-AUC: 0.7017434119508382
Classification Report:

```

	precision	recall	f1-score	support
0	0.80	1.00	0.89	7917
1	0.84	0.02	0.04	2083
accuracy			0.80	10000
macro avg	0.82	0.51	0.46	10000
weighted avg	0.80	0.80	0.71	10000

## Metriche:

- Accuracy: buona
- ROC-AUC: moderato
- Recall per classe 'furto': bassa (problema di sbilanciamento)

06



# Conclusioni

## Il modello aiuta a:

- Evidenziare pattern temporali e spaziali.
- Supportare le forze dell'ordine nella pianificazione.

## Miglioramenti

- Modelli più complessi (XGBoost, reti neurali).
- Aggiornamento con dati più recenti.





# Grazie!

**Yanik Veselinov Dimitrov**

---