

# A Caccia di Fake News

## PCTO - Summer Camp Informatica & Intelligenza Artificiale

Giovedì, 10 Settembre 2026

Michael Soprano - [michael.soprano@uniud.it](mailto:michael.soprano@uniud.it)

Università degli Studi di Udine



**UNIVERSITÀ  
DEGLI STUDI  
DI UDINE**

HIC SUNT FUTURA

# Scaletta

1. Riconoscere il problema
2. Come si verifica una notizia?
3. Può aiutarci l'intelligenza artificiale?
4. Mettiamo alla prova un modello

## Michael Soprano

**Ricercatore** @ Università degli Studi di Udine

Dipartimento di Scienze Matematiche, Informatiche e Fisiche (DMIF)

[Laboratorio Social, Mobile, Data & Crowd \(SMDC\)](#)

Mi occupo di **disinformazione, fact-checking, crowdsourcing e intelligenza artificiale**



Questa immagine è autentica?





# Un'immagine manipolata

- Nell'originale, **Messi non pesta il piede di Spence**



*Immagine originale*



*Immagine manipolata*

# Fake news

- Espressione molto diffusa, ma **ampia e poco precisa**
- Può riferirsi a contenuti **falsi o fuorvianti** presentati come informazione attendibile
- Può raggruppare fenomeni diversi, che è utile distinguere in modo più preciso

# Tipologie di informazione

- **Misinformazione**

- Informazione falsa o fuorviante condivisa **senza intenzione di causare un danno**

- **Disinformazione**

- Informazione falsa o fuorviante creata o diffusa **intenzionalmente per causare un danno**

- **Malinformazione**

- Informazione basata su fatti reali, ma utilizzata **per causare un danno**

## Fact-checking

- Processo di **verifica di un'affermazione** sulla base delle evidenze disponibili
- Tre elementi sono fondamentali:
  - **Affermazione**: che cosa viene sostenuto?
  - **Evidenze**: quali informazioni pertinenti e affidabili abbiamo?
  - **Verdetto**: che cosa possiamo concludere?
- Il verdetto deve essere **motivato dalle evidenze**



# Un caso di fact-checking

- Consideriamo questa affermazione:

“*Anche grazie alle misure del Governo [...] abbiamo drasticamente ridotto il tasso di abbandono scolastico. Nel 2022 era all'11,5%, oggi è al 7,3%*”

- Per verificarla dobbiamo chiederci:
  - Da dove provengono i **numeri**?
  - Sono dati **osservati o stimati**?
  - Qual è il dato ufficiale più recente?
  - I dati dimostrano un rapporto di **causa ed effetto**?

# Analisi delle evidenze

- Il calo dell'abbandono scolastico è **reale**
  - 11,5% nel 2022
  - **8,2% nel 2025**, ultimo dato ufficiale ISTAT
- Per il 2026, INVALSI stima che il tasso **potrebbe scendere al 7,3%**
  - È una **previsione**, non ancora un dato ufficiale
- La diminuzione era iniziata **prima del 2022**
- I dati, da soli, non permettono di stabilire quanto il calo dipenda dalle **misure del governo**

# Le verifiche come dati

- Una raccolta di verifiche può essere organizzata come un **dataset**
- Ogni esempio può associare:
  - Un'**affermazione**
  - Le **evidenze** utilizzate
  - Un'**etichetta di riferimento**
- Le etichette di riferimento costituiscono la **ground truth**
- Questi dati possono essere utilizzati per **addestrare** o **valutare** sistemi automatici

## Automatizzare il fact-checking

- La quantità di contenuti da verificare rende difficile affidarsi soltanto a **controlli manuali**
- Un sistema automatico può intervenire in diverse fasi:
  - Individuare le affermazioni **da verificare**
  - Cercare e selezionare **evidenze pertinenti**
  - Confrontare **affermazioni ed evidenze**
  - Contribuire alla formulazione di un **verdetto**
- Nel laboratorio ci concentreremo su **un solo passaggio**

# Il compito di classificazione

- Forniremo al modello:
  - Un'**affermazione**
  - Un'**evidenza già selezionata**
- Formuleremo il problema come **Natural Language Inference (NLI)**
- Il modello dovrà stabilire se l'evidenza:
  - **Conferma** l'affermazione → Entailment
  - La **contraddice** → Contradiction
  - **Non è sufficiente** → Neutral
- Evidenza insufficiente *non significa che l'affermazione è falsa*

# Esercizio di classificazione

- **Affermazione**

- Nel 2025 l'Italia ha avuto la migliore crescita delle esportazioni verso gli Stati Uniti tra tutti i Paesi dell'Unione europea

- **Evidenza**

- Nel 2025 le esportazioni italiane verso gli Stati Uniti sono cresciute del **7,2%** rispetto al 2024

- **Quale classe assegnereste?**

- Entailment
- Contradiction
- Neutral



# Il ruolo dell'evidenza

- Consideriamo sempre la stessa affermazione, ma cambiamo l'**evidenza disponibile**
- Nell'ultimo caso l'evidenza sarebbe **falsa**, ma supporterebbe comunque l'affermazione
- Il modello valuta la **relazione tra i testi**, non verifica automaticamente l'affidabilità dell'evidenza

| Evidenza fornita                                                                                       | Classe        |
|--------------------------------------------------------------------------------------------------------|---------------|
| Nel 2025 le esportazioni italiane verso gli USA sono cresciute del <b>7,2%</b>                         | Neutral       |
| Nel 2025 le esportazioni irlandesi verso gli USA sono cresciute del <b>52%</b>                         | Contradiction |
| Nel 2025 nessun altro Paese UE ha superato il <b>7,2%</b> di crescita delle esportazioni verso gli USA | Entailment    |

## Procedura di valutazione

1. Partiamo da esempi con una **ground truth**
2. Il modello riceve **affermazione + evidenza**
3. Produce una **previsione**
4. Confrontiamo previsione e ground truth
5. Misuriamo quanto spesso il modello **classifica correttamente**

# Strumenti utilizzati

- Useremo `mDeBERTa-v3-base-xnli-multilingual-nli-2mil7`
  - Modello multilingue per la **Natural Language Inference**
  - Può lavorare anche con testi in **italiano**
- Eseguiremo il codice su **Google Colab**
  - Direttamente dal browser
  - Senza installare Python
  - Con GPU, se disponibile
- Aprite il notebook e create una **copia personale**
  - Link: <https://bit.ly/3T0cRH5>

# Dataset di prova

- Il notebook contiene **12 casi costruiti a partire da verifiche reali**, quattro per classe
- Ogni esempio contiene:
  - Un'**affermazione**
  - Un'**evidenza**
  - La **ground truth**
  - Il link alla **fonte**
- La ground truth è la classe di riferimento con cui confronteremo il modello

| Caso                                                          | Ground truth  |
|---------------------------------------------------------------|---------------|
| Sinner è il primo italiano n. 1 ATP                           | Entailment    |
| Messi pesta il piede di Spence                                | Contradiction |
| Le politiche del governo hanno causato il calo dell'abbandono | Neutral       |

# Preparazione del notebook

- **Eseguite le celle** dall'alto verso il basso
- Le prime celle:
  - Installano le librerie
  - Caricano dataset e modello

```
from transformers import pipeline

classifier = pipeline(
    "text-classification",
    model=MODEL_NAME,
)
```

# Analisi di un caso reale

- **Affermazione**

- Messi ha pestato il piede di Spence nell'azione del secondo gol argentino

- **Evidenza**

- Foto originale e video mostrano che Messi **salta il difensore senza pestargli il piede**

- Prima di eseguire:

- Quale classe vi aspettate?

```
prediction, scores = classify(  
    evidence,  
    claim,  
)  
  
print(prediction)  
print(scores)
```

# Interpretazione dei punteggi

Un possibile output è:

```
{  
  "Entailment": 0.01,  
  "Neutral": 0.03,  
  "Contradiction": 0.96  
}
```

- Il punteggio più alto determina la **previsione**
- 0.96 non significa che l'affermazione sia «falsa al 96%»
- Esprime la **preferenza del modello per quella classe rispetto alle altre**

# Valutazione del modello

- Eseguiamo il modello su **tutti i 12 esempi**
- Per ciascuno confrontiamo:
  - `ground_truth`
  - previsione
- L'**accuratezza** è la percentuale di esempi classificati correttamente
- Guardiamo anche **quali casi vengono sbagliati**

```
accuracy = results["correct"].mean()  
  
print(f"Accuratezza: {accuracy:.1%}")
```



# Interpretazione dell'accuratezza

- Il dataset contiene soltanto **12 esempi**
- I casi sono stati scelti per questa **esercitazione**
- Il risultato descrive quindi il comportamento del modello **su questi esempi**
- Per una valutazione reale servirebbe un dataset:
  - Molto più ampio e vario
  - Costruito indipendentemente dal modello

# Esercitazione

- Scegliete **alcuni casi** del dataset
- Prima di eseguire il modello:
  - Leggete affermazione ed evidenza
  - Formulate la vostra **previsione**
- Confrontatela poi con quella del modello
- Provate almeno un caso per classe:
  - Entailment
  - Contradiction
  - Neutral

# Analisi di robustezza

- Partite da un esempio e cambiate **un elemento alla volta**
- Provate a modificare:
  - Una negazione
  - Una data o un numero
  - Una parola importante
  - Una parte dell'evidenza
- La previsione cambia?
- Riuscite a **far sbagliare il modello o a renderlo incerto?**

# Nuovi casi di prova

- Scegliete un nuovo caso da:
  - [Facta](#)
  - [Pagella Politica](#)
- Individuate:
  - L'**affermazione**
  - Una breve **evidenza pertinente**
- Stabilite la classe attesa
- Provate infine la coppia con il modello

# Limiti dell'approccio

- Nel laboratorio l'evidenza è **già stata selezionata**
- Un sistema completo dovrebbe anche:
  - Trovare l'affermazione da verificare
  - Cercare fonti **affidabili**
  - Selezionare le evidenze pertinenti
  - Valutarne qualità e contesto
- La NLI è quindi **solo una parte** del processo di **fact-checking**

# Conclusioni

- Un contenuto plausibile può essere **falso o manipolato**
- Verificare un'affermazione richiede **fonti ed evidenze**
- La conclusione dipende da **quale evidenza abbiamo a disposizione**
- Possiamo valutare un modello confrontandolo con una **ground truth**
- L'intelligenza artificiale può supportare il fact-checking, ma **può sbagliare**

# Risorse e approfondimenti

- Fact-checking in Italia: [Pagella Politica](#)
- Disinformazione e verifica dei contenuti: [Facta](#)
- Modelli e dataset: [Hugging Face](#)
- Ricerca di fact-check: [Google Fact Check Explorer](#)