



**XIII**

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

# Sinistri causati da eventi naturali: come identificarli grazie al machine learning

Cristiana Di Russo

Attuariato e Pricing – Direzione Danni

Groupama Assicurazioni



# Agenda

---

**XIII**

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

- I. Contesto e Obiettivo
- II. Metodologia
- III. Applicazione



# Contesto e Obiettivo

## XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

## Contesto

- Negli ultimi anni si è osservato un aumento della frequenza e dell'intensità di eventi naturali.
- Gli eventi estremi hanno forte impatto sul territorio e sulla popolazione.
- Diversi processi delle compagnie assicurative possono beneficiare di una corretta identificazione dei sinistri causati da eventi naturali.

## Obiettivo

- Definire uno strumento analitico per il riconoscimento, la classificazione ed il raggruppamento dei sinistri causati da eventi naturali.
- Come? Attraverso tecniche di Data Mining e Machine Learning.



# Metodologia

# XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

- a) Identificare i sinistri riconducibili ad un **evento atmosferico/naturale noto**, ovvero di cui si conoscono data e coordinate geografiche.
  
- b) **Raggruppare i sinistri** avvenuti nello stesso periodo temporale in zone limitrofe e determinare se sono stati causati da uno stesso evento atmosferico/naturale.



# Caso A: Evento Naturale Noto

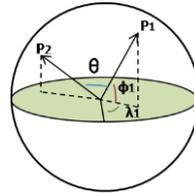
# XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

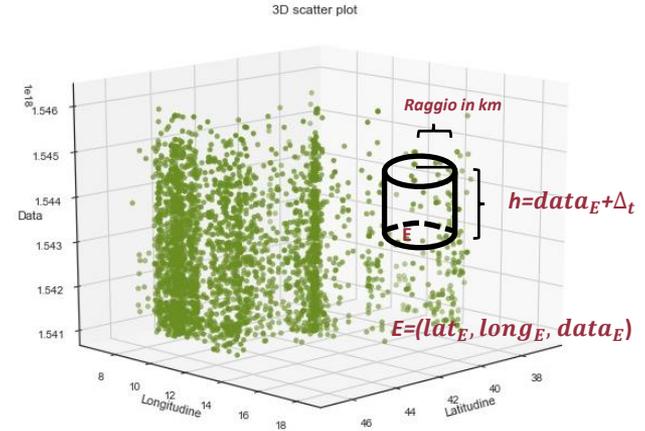
- A partire da un evento noto, identificazione dei sinistri avvenuti in un intervallo temporale ed in zone limitrofe all'evento.
- **Parametri:** data, latitudine, longitudine dell'evento, delta temporale e spaziale della ricerca.
- Distanza geodetica tra due punti data la loro latitudine e longitudine.



$$d(P_1, P_2) = 2 \cdot r \cdot \arcsin \sqrt{\sin^2 \left( \frac{\phi_1 - \phi_2}{2} \right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2 \left( \frac{\lambda_1 - \lambda_2}{2} \right)}$$

dove  $r$  è il raggio della sfera,  $\lambda_i$  e  $\phi_i$  sono rispettivamente la longitudine e la latitudine del punto  $P_i$ .

- **Output:** sinistri che soddisfano le condizioni temporali e spaziali passate in input.





# Applicazione Caso A

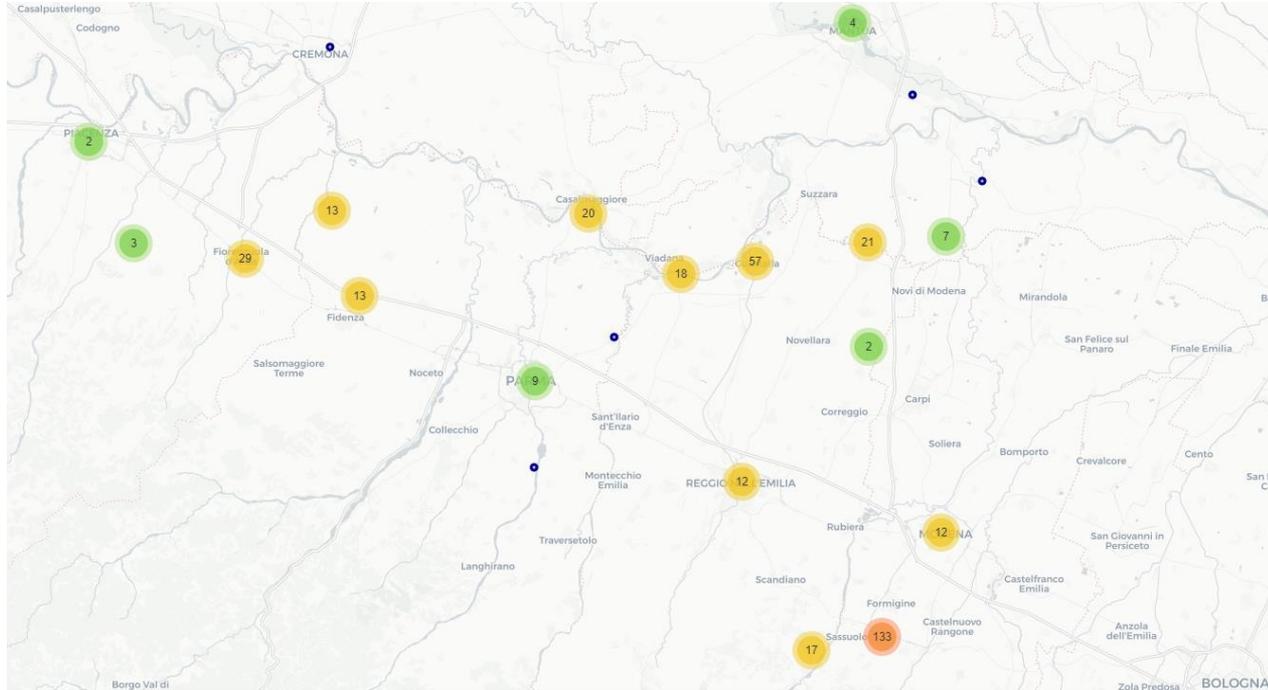
# XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

## Evento 26 luglio 2021 provincia di Parma





## Caso B: Evento Naturale Non Noto

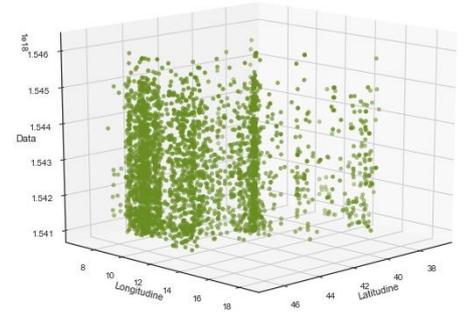
XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

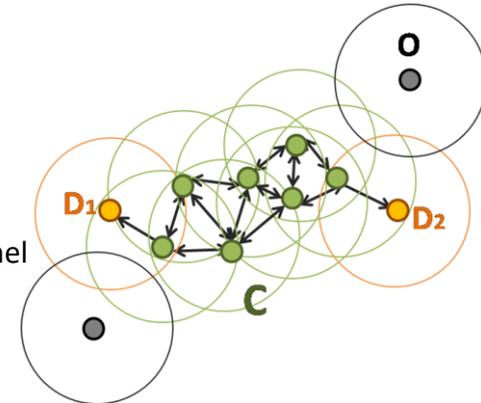
ROMA  
10-12 Novembre 2021

- Identificare e raggruppare i sinistri avvenuti nello stesso periodo temporale in zone limitrofe.
- Machine Learning non supervisionato: metodo di clustering basato sulla densità dei punti presenti in una regione.



### DBSCAN (Density-Based Spatial Clustering of Applications with Noise [1])

- Parametri
  - $MinPts$ : numero minimo di punti per formare un cluster,
  - $\epsilon$ : soglia per la distanza.
- L'algoritmo classifica i punti come:
  - «Core point»: almeno  $MinPts$  punti appartengono al suo intorno di raggio  $\epsilon$ .
  - «Border point»: non verifica il criterio del numero minimo di punti presenti nel suo intorno di raggio  $\epsilon$  ma, appartiene all'intorno di un core point.
  - «Outlier»: non appartiene alle due categorie precedenti.





## Caso B: Evento Naturale Non Noto

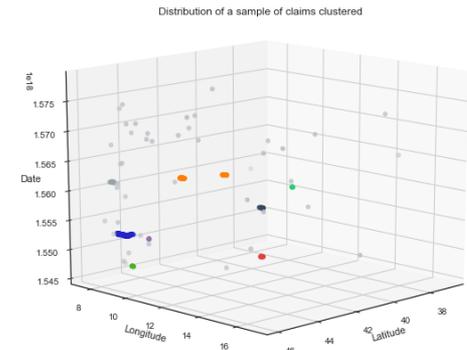
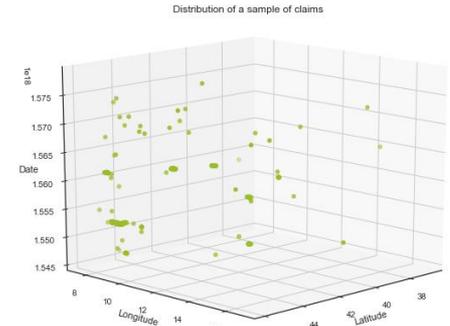
XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

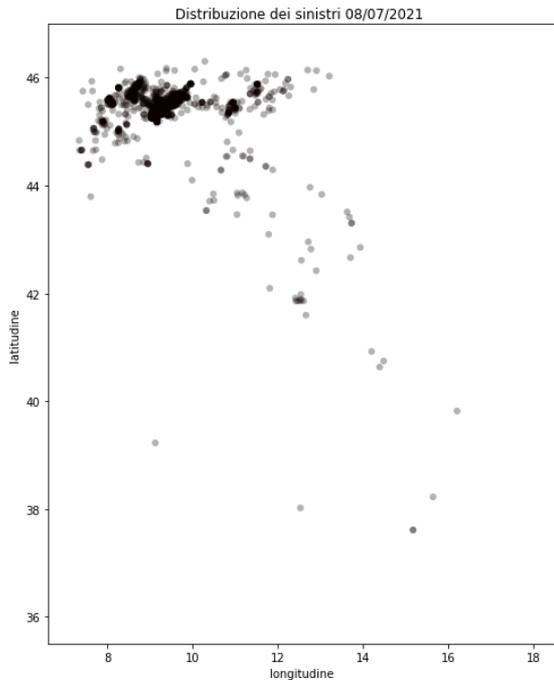
- Gestione di dati spazio-temporali.
- Generalizzazione dell'algoritmo: **ST-DBSCAN [2]** prende in considerazione sia la dimensione spaziale che temporale.
- **Parametri**
  - $MinPts$ : numero minimo di punti per formare un cluster,
  - $\varepsilon_S$ : soglia distanza spaziale (misurata in km),
  - $\varepsilon_T$ : soglia distanza temporale (misurata in giorni).
- Output: sinistri raggruppati secondo i parametri spaziali e temporali passati in input.



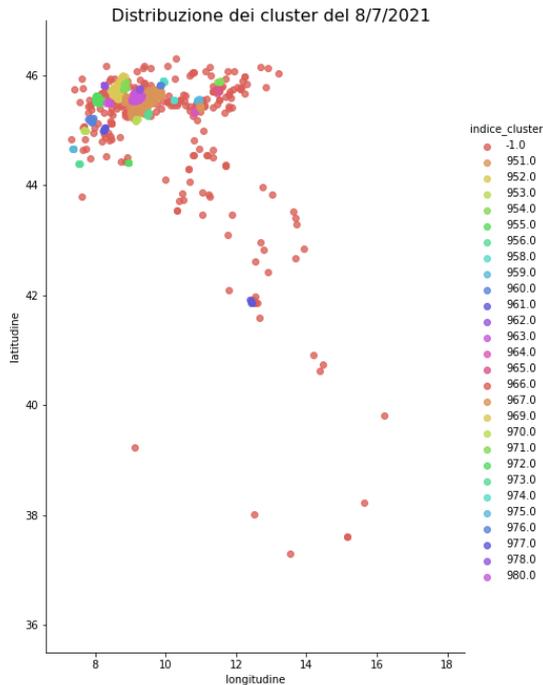


# Applicazione Caso B

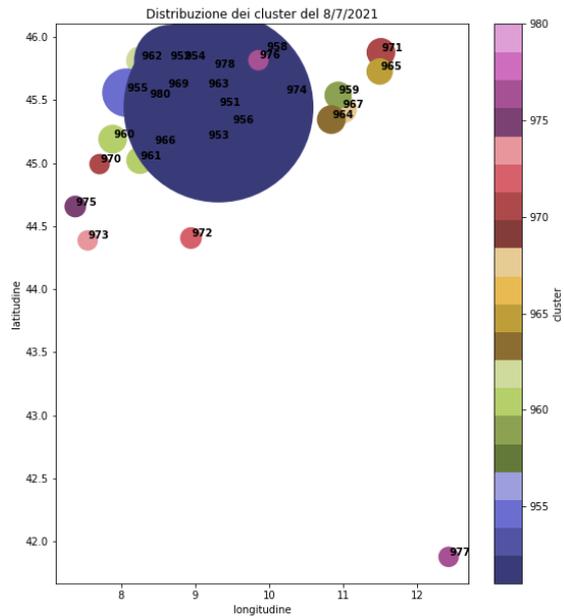
Il grafico mostra la distribuzione dei sinistri avvenuti in un giorno specifico.



Il grafico mostra la distribuzione dei cluster e degli outlier (i sinistri che non appartengono a nessun cluster) di un giorno specifico.



Il grafico rappresenta la distribuzione dei cluster di un giorno specifico.



# XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021



# Applicazione Caso B

# XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

Grafico 2D che mostra la distribuzione dei cluster e degli outlier (ovvero i sinistri che non appartengono a nessun cluster) associati ad un evento specifico avvenuto in più giorni.

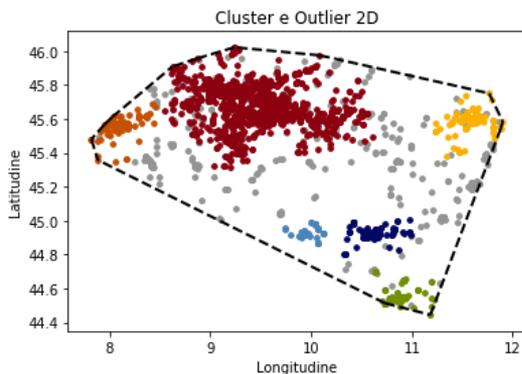
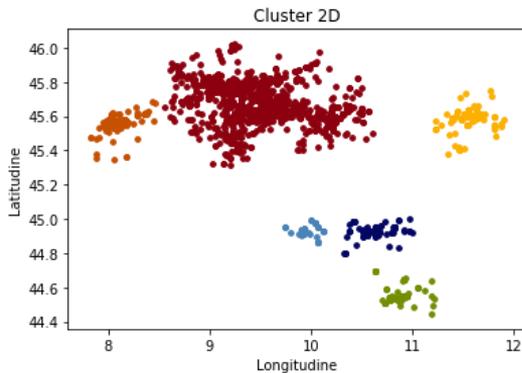
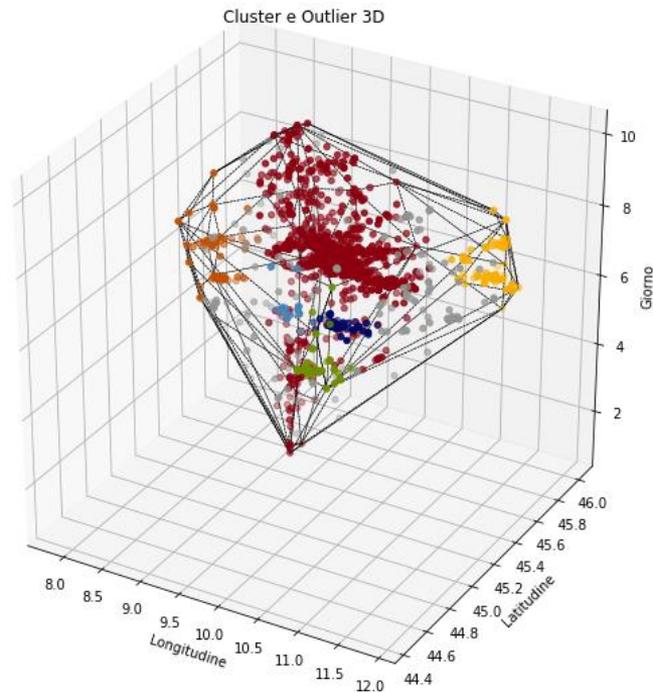


Grafico 3D che mostra la distribuzione dei cluster e degli outlier (i sinistri che non appartengono a nessun cluster) associati ad un evento specifico avvenuto in più giorni.





# Bibliografia

# XIII

CONGRESSO  
NAZIONALE  
DEGLI  
ATTUARI

INNOVAZIONE  
TECNOLOGICA  
E RISCHI SISTEMICI:  
L'ATTUARIO  
VALUTATORE  
GLOBALE  
DELL'INCERTEZZA

ROMA  
10-12 Novembre 2021

- 1) Ester Martin, Kriegel Hans-Peter, Sander Jorg, Xu Xiaowei, *A density-based algorithm for discovering clusters in large spatial databases with noise*. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (1996).
- 2) Birant Derya, Kut Alp, *ST-DBSCAN: An algorithm for clustering spatial-temporal data*. Data & Knowledge Engineering, 60 (2007).

## Grazie per l'attenzione!



### Confidentiality Statement

Questa presentazione è stata preparata per il XIII Congresso Nazionale degli Attuari ed è da considerarsi riservata per tale uso.