1. ABSTRACT

Con la diffusione di un tenore di vita superiore, la popolazione mondiale ha visto incrementare notevolmente la sua aspettativa di vita negli ultimi decenni. Se da un lato vivere più a lungo rappresenta indiscutibilmente un fattore positivo, dall’altro si assiste ad un invecchiamento della popolazione. Questo porta con sé numerose sfide da affrontare perché l’obiettivo di ognuno è vivere ben, a lungo.

Una delle patologie che presentano il maggior tasso di incidenza nelle chiamate ricevute da AREU (Agenzia Regionale Emergenze Urgenze), ossia l’agenzia che si occupa della gestione del numero telefonico riservato alle emergenze mediche, è l’ictus. Incide infatti per il 33.9% degli interventi e conta più di 9000 casi all’anno, in costante aumento.

La popolazione colpita ha mediamente 74 anni per quanto riguarda gli individui maschi e 79 anni per le femmine. Sono proprio questi a suggerire una possibile correlazione con l’età del paziente e considerato il fenomeno aging, indagare meglio sulle possibili cause che portano alla formazione di un ictus appare fondamentale in ottica futura. Nei prossimi 20 anni è infatti previsto un raddoppio dei casi per via dell’invecchiamento della popolazione.

L’obiettivo di questo progetto di tesi è quello di individuare possibili correlazioni che leghino gli eventi di ictus a fenomeni meteorologici e agli agenti inquinanti, che come suggerito dalla letteratura (elenco paper), sembrano avere effetti non trascurabili. L’obiettivo finale è quello di creare una funzione di rischio che possa predire, in una certa misura, il possibile insorgere della patologia.

1. INTRODUCTION
   1. WHAT A STROKE IS
   2. RISK FACTORS
   3. RESCUE PROTOCOL

Il protocollo di soccorso prevede che un operatore del 118 riceva una chiamata per una richiesta di intervento. Per individuare un sospetto caso di ictus si fa uso della CPSS (Cincinnati Prehospital Stroke Scale), ossia una metrica per valutare la condizione del paziente chiedendo:

* Parla male?
* Ha la bocca storta?
* Riesce a tenere le braccia sollevate?
* Da quanto tempo ha questo problema?

Tali domande mirano ad individuare rispettivamente:

* Anomalie del linguaggio
* Paresi facciale
* Deficit motorio degli arti superiori

Sulla base delle risposte ai precedenti quesiti, gli operatori attribuiscono un codice di gravità all’intervento e dispongono l’invio dei mezzi più adeguati. Tra questi mezzi figurano:

* Mezzo di soccorso di base (MSB): un’ambulanza con a bordo personale tecnico soccorritore.
* Mezzo di soccorso intermedio (MSI): un’ambulanza che preveda a bordo un infermiere attrezzato e autorizzato ad operare procedure di intervento avanzato.
* Mezzo di soccorso avanzato (MSA): un mezzo su gomma (ambulanza o automedica) e a pala (elicottero) con a bordo un infermiere ed un medico di pronto soccorso.

Giunti sul posto, gli operatori del 118 verificano le condizioni del paziente ed eventualmente attribuiscono alla missione il codice ictus, ossia se la CPSS risulta positiva, il paziente ha più di 18 anni e i sintomi presenti da non oltre 4 ore. Il tal caso gli operatori comunicano con l’hub centrale al fine di farsi indirizzare verso la struttura ospedaliera più adeguata al trattamento. Altrimenti si procede al trasporto presso il pronto soccorso più vicino.

Gli ospedali attrezzati al trattamento di un ictus, ossia quelli dotati di stroke unit (UCV) sono classificati in due livelli:

* Livello II: è disponibile il solo trattamento fibrinolitico per lo scioglimento del trombo.
* Livello III: sono disponibili una sala attrezzata e del personale specializzato nell’asportazione meccanica del trombo.

E’ compito dell’hub del 118 inviare l’allerta circa una paziente in arrivo all’ospedale di destinazione, che procederà ad allertare il neurologo di riferimento. Il neurologo svolge il ruolo di attending doctor.

Dopo l’arrivo in pronto soccorso, il paziente viene immediatamente inviato in radiologia sottoposto ad una TAC per verificare lo stato cerebrale. Successivamente si procede con la terapia fibrinolitica o con l’asportazione meccanica del trombo.

* 1. STROKE ORIGIN AND TREATMENT

Il trattamento di un caso di ictus prevede la dissoluzione o la rimozione meccanica di un trombo. Il trombo è costituito da una massa ematica solida che si origina all’interno del sistema cardiocircolatorio. Generalmente si tratta di un coagulo di sangue costituito da globuli rossi, bianchi, piastrine e fibrina.

* Trattamento fibrinolitico: l’obiettivo è quello di disgregare la fibrina che tiene unito il coagulo in modo da ripristinare la normale circolazione
* Asportazione meccanica: tramite intervento chirurgico si individua il trombo e si procede alla sua asportazione mediante l’uso di un catetere endovenoso.
  1. MEDICAL CONSEQUENCES AND IMPACT ON THE SOCIAL SECURITY SYSTEM

1. STROKE: WEATHER CONDITIONS AND POLLUTANTS
   1. THE STATE OF THE ART

Il tentativo di correlare gli ictus agli eventi meteorologici e agli agenti inquinanti non è totalmente nuovo in letteratura. Esistono infatti diverse pubblicazioni che trattano l’argomento seppur le conclusioni siano talvolta discordanti.

In tabella sono riportate le pubblicazioni analizzate:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| TITOLO | AUTORI | ANNO | | POPOLAZIONE ANALIZZATA | | PARAMETRI MONITORATI | CONCLUSIONI |
| Air Pollution and Hospital Admissions for Ischemic and Hemorrhagic Stroke Among Medicare Beneficiaries | Gregory A. Wellenius, Joel Schwartz, Murray A. Mittleman | 2005 |  | |  | |  |
| Air Pollution  A New Risk Factor in Ischemic Stroke Mortality | Yun-Chul Hong, Jong-Tae Lee, Ho Kim, Ho-Jang Kwon | 2002 |  | |  | |  |
| Air Pollution and Stroke | Kuan Ken Lee, Mark R. Miller, Anoop S. V. Shah | 2018 |  | |  | |  |
| Ambient Air Pollution and Risk for Ischemic Stroke:  A Short-Term Exposure Assessment in South China | Pi Guo, Yulin Wang, Wenru Feng, Jiagang Wu, Chuanxi Fu, Hai Deng, Jun Huang,  Li Wang, Murui Zheng, Huazhang Liu | 2017 |  | |  | |  |
| Evidence for an Association Between Air Pollution and daily Stroke Admissions in Kaohsiung, Taiwan | Shang-Shyue Tsai, William B. Goggins, Hui-Fen Chiu, Chun-Yuh Yang | 2003 |  | |  | |  |
| Ambient Temperature and Stroke Risk  Evidence Supporting a Short-Term Effect at a Population Level from Acute Environmental Exposures | Pablo M. Lavados, Verónica V. Olavarría, Lorena Hoffmeister | 2017 |  | |  | |  |

* 1. TEMPERATURE CORRELATION
  2. POLLUTANTS CORRELATION

1. THE DATA
   1. THE DATABASE OF STOKES

I dati utilizzati per lo svolgimento di questo progetto di tesi sono stati forniti da AREU (Agenzia Regionale Emergenze Urgenze), ossia l’agenzia che si occupa della gestione delle chiamate di emergenza a livello regionale in Lombardia.

AREU è infatti in grado di raccogliere dati all’interno del proprio sistema informativo a partire dalla ricezione di una chiamata al 118. Queste informazioni comprendono i dati del paziente, gli orari di dispatch dei mezzi, le coordinate del luogo dell’evento e una serie di altre informazioni meglio descritte in seguito.

I dati fornitimi sono stati anonimizzati e rappresentano un sottoinsieme di quanto presente nel sistema informativo AREU.

* + 1. THE INFORMATION SYSTEM ARCHITECTURE

La base dati detenuta da AREU trova posto all’interno di un sistema informativo ben più ampio.

L’architettura dello stesso è basata su quattro entità fondamentali:

* **Evento:** è un’entità atomica alla quale vengono ricondotte una o più chiamate al 118
* **Missione:** per ogni evento viene avviata una missione, alla quale farà riferimento ogni mezzo coinvolto nell’intervento
* **Tratta:** rappresenta un frammento del tragitto di soccorso che separa la centrale dal paziente e il paziente dall’ospedale e monitora i tempi di interarrivo.
* **Paziente:** contiene una descrizione anagrafica del paziente e del suo stato di salute pregresso

Alcuni dei campi presenti all’interno del database risultano essere duplicati. La causa è da ricercarsi nel normale processo evolutivo a cui le sorgenti dati vanno in contro nel tempo ed è da associare alla necessità di raccogliere maggiori informazioni sia sui pazienti sia sull’efficienza interna dei processi aziendali, al fine di migliorare la prestazione erogata. Alcuni campi invece, quali ad esempio i tempi di interarrivo, sono stati materializzati al fine di ottimizzare le operazioni di interrogazione a scopo di business intelligence.

* + 1. DATA ACQUISITION METHODOLOGY

I dati inseriti all’ìnterno del sistema informativo AREU provengono da una molteplicità di fonti. Derivano infatti da un processo di raccolta realizzato durante le due principali fasi dell’intervento:

* Chiamata al 118
* Intervento sul posto dei sanitari

Lo scopo della chiamata al 118 è quello di raccogliere i sintomi del paziente, accertare l’indirizzo ed erogare un codice di intervento.

Spetta poi ai sanitari che si recano sul luogo verificare l’anagrafica, raccogliere la posizione GPS, accertare patologie pregresse ed eventualmente correggere il codice di intervento.

I dati sugli intertempi sono calcolati in modo automatico e vengono impiegati per lo più a livello statistico.

* + 1. DESCRIPTION OF THE FIELDS OF THE DATABASE

|  |  |  |
| --- | --- | --- |
| VOCE | TIPO | DESCRIZIONE |
| AAT | String | Articolazioni Aziendali Territoriali |
| ANNO | Integer | Year |
| MESE | Integer | Month |
| ID\_EMERGENCY | ID | ID of the event |
| ID\_EMERG\_HOUR\_IN\_DAY | Integer | Hour of the event |
| CD\_SEVERITY\_LEVEL | Code | Severity |
| CD\_EMERG\_CLASS | Code | Classification |
| CD\_EMERG\_CLASS\_DTL | Code | Detailed classification |
| CD\_REASON | Code | Reason |
| CD\_REASON\_DTL | Code | Detailed reason |
| CD\_CALLER | Code | Caller |
| CD\_PLACE | Code | Place |
| CD\_PROVINCE\_ISTAT | Code | Province ISTAT code |
| CD\_TOWN\_ISTAT | Code | Town ISTAT code |
| DS\_TOWN | Description | Town name |
| CD\_PROVINCE | Code | Province code |
| CD\_STREET | Code | Street code |
| DS\_STREET | Description | Street |
| DS\_STREET\_TYP | Description | Street type |
| CD\_TOWN\_ISTAT\_STREET | Code | Street ISTAT code |
| CD\_ZONE | Code | Zone code |
| CD\_INSTITUT\_TRANSP\_TO | Code | Code of the zone the patient is transported to |
| CD\_INSTITUT\_TRANSP\_FROM | Code | Code of the zone the patient is transported from |
| CD\_METEO | Code | Weather |
| ID\_EMERG\_MONTH | Date-Time | YearMonth of the event |
| DT\_EMERG\_DAY | Date-Time | Day of the emergency |
| DT\_EMERG\_OPEN | Date-Time | Open time for the event |
| DT\_EMRG\_OPEN\_HH24MI | Date-Time | Open hour for the event |
| DT\_EMERG\_CLOSE | Date-Time | Close time for the event |
| DT\_MISSION\_OPEN\_1 | Date-Time | Open time for mission 1 |
| DT\_CAR\_START\_1 | Date-Time | First vehicle departure |
| DT\_CAR\_H\_1 | Date-Time | First vehicle reaching the hospital |
| DT\_CAR\_I\_1 | Date-Time | First vechicle reaching the emergency |
| VL\_PEOPLE | Integer | Involved people |
| VL\_PATIENT | Integer | Patients |
| VL\_PATIENT\_ECG | Integer | Patient did ECG |
| VL\_EMERG\_INTERVAL | Integer | Event duration |
| VL\_MISSION\_OPEN\_1\_INTERVAL | Time | First vehicle alerted |
| VL\_CAR\_START\_1\_INTERVAL | Time | Fisrt vechicle departure |
| VL\_CAR\_H\_1\_INTERVAL | Time | First vehicle reaches the hospital |
| VL\_CAR\_H\_1\_INTERVAL\_HH24\_MI | Time | Time required to reach the hospital |
| VL\_CAR\_H\_1\_INT\_MINUTI | Time | Time required to reach the hospital in minutes |
| VL\_CAR\_I\_1\_INTERVAL | Time | First vehicle reaches the patient |
| VL\_CAR\_I\_1\_INT\_HH24\_MI | Time | Time required to reach the patient |
| VL\_CAR\_I\_1\_INT\_MINUTI | Time | Time required to reach the patient in minutes |
| VL\_AMOUNT\_MISSION\_MSB | Integer | Number of MSB involved |
| VL\_AMOUNT\_MISSION\_MSI | Integer | Number of MSI involved |
| VL\_AMOUNT\_MISSION\_MSA | Integer | Number of MSA involved |
| DT\_CAR\_I\_1\_MSB | Date-Time | MSB reaching the patient |
| DT\_CAR\_I\_1\_MSI | Date-Time | MSI reaching the patient |
| DT\_CAR\_I\_1\_MSA | Date-Time | MSA reaching the patient |
| VL\_CAR\_I\_1\_INTERVAL\_MSB | Integer | Time required by the MSB to reach the patient |
| VL\_CAR\_I\_1\_INTERVAL\_MSI | Integer | Time required by the MSI to reach the patient |
| VL\_CAR\_I\_1\_INTERVAL\_MSA | Integer | Time required by the MSA to reach the patient |
| VL\_GEO\_X | Coordinate | EAST coordinate of the event |
| VL\_GEO\_Y | Coordinate | NORTH coordinate of the event |
| AC\_SESSO | Categorical | Sex |
| VL\_ETA | Integer | Age |
| DS\_PAT\_1 | Description | Known pathology1 for the patient |
| DS\_PAT\_2 | Description | Known pathology2 for the patient |
| ID\_CODICE | Code | Event severity code |
| ID\_CODICE\_E | Code | Filtered severity code |
| OSPEDALE\_DENOMINAZIONE | String | Destination hospital |
| OSPEDALE\_LOCALITA | String | Hospital town |
| OSPEDALE\_REPARTO | String | Hospital ward |
| OSPEDALE\_CODICE | Code | Hospital code |
| OSPEDALE\_DESCR\_SINTETICA | String | Hospital synthetic description |
| OSPEDALE\_DESCR\_COMPLETA | String | Hospital complete description |
| OSPEDALE\_PROVINCIA | String | Hospital province |
| DS\_REPARTO\_R | Descrizione | Ward description |
| DT\_APERTURA | Date-Time | Mission opening time |
| ID\_CODICE\_R | Code | Return code |
| ID\_CODICE\_E2 | Code | Emergency code |
| ID\_PAT1\_LIV1 | Code | Known pathology1 level1 code |
| ID\_PAT1\_LIV2 | Code | Known pathology1 level2 code |
| ID\_PAT2\_LIV1 | Code | Known pathology2 level1 code |
| ID\_PAT2\_LIV2 | Code | Known pathology2 level2 code |
| DT\_COMPARSA\_EV | Date-Time | Sympthoms appearance time |
| DT\_REGISTRAZ\_EV | Date-Time | Event registration time |
| VL\_CAR\_H\_1\_COMPARSA\_EV\_MIN | Time | Time between sympthoms and vehicle arrival |
| codice ictus | Code | Stroke code |

* 1. WEATHER STATIONS DATASET
     1. STATIONS DISTRIBUTION AND DATA
     2. MEASURED VALUES AND SAMPLING INTERVAL
     3. DATABASE STRUCTURE
  2. AIR QUALITY STATIONS
     1. STATIONS DISTRIBUTION AND DATA
     2. MEASURED VALUES AND SAMPLING INTERVAL
     3. DATABASE TRUCTURE
  3. DEMOGRAPHIC DATA
     1. RESIDENT POPULATION FOR EVERY PROVINCE

1. DATA EXPLORATION
   1. STROKES DISTRIBUTION DURING THE WEEK
   2. STROKES DISTRIBUTIONS DURING THE YEAR
   3. STROKES DISTRIBUTION EVER THE TERRITORY
      1. DEMOGRAPHIC INCIDENCE
      2. GIS
   4. STROKES PERIODICITY
   5. POSSIBLE CORRELATIONS
2. ANALYSIS
   1. DATA MANIPULATION ADN METHODOLOGIES
      1. DATA MANAGEMENT AND USAGE
      2. UNAVAILABLE DATA DELETION
   2. THE ANALYSIS PROCESS
      1. TREND ANALYSIS
      2. LA PERIODS ANALYSIS
   3. THE RISK FUNCTION
3. DATA INTERPOLATION
   1. THE STATE OF THE ART
      1. ARPA OPTIMAL INTERPOLATION

Lo stato dell’arte per quanto concerne lo studio, l’analisi le interpolazioni dei dati ambientali in Italia è rappresentato dal sistema ARPA (Agenzia Regionale per la Protezione dell’Ambiente). ARPA si occupa di raccogliere dati riguardanti sia i fenomeni meteorologici sia gli agenti inquinanti in tutta la penisola; questa raccolta avviene mediante l’impiego di stazioni fisse e mobili sparse su tutto il territorio.

Questi dati concorrono alla creazione di mappe e previsioni consultabili direttamente direttamente sul portale web di ARPA.

Poiché i dati sono raccolti dalle centraline in maniera puntuale, è di fondamentale importanza interpolare correttamente questi valori al fine di creare una mappa che copra l’intero territorio di interesse con valori quanto più accurati possibile.

Il modello di adottato da ARPA consiste in una interpolazione spaziale tridimensionale che prende il nome di Optimal Interpolation. I parametri forniti in ingresso al modello sono le coordinate delle centraline e le loro altitudini.

* + 1. CIRCUMNSTANTIAL DIFFERENCES AND DEFINITION OF A DIFFERENT MODEL

Nonostante ARPA disponga già di un modello di interpolazione efficace, gli algoritmi da loro utilizzati fanno uso di librerie personalizzate e soggette a licenza d’uso, nonché ottimizzate appositamente per l’hardware su cui il software viene eseguito. Per questo motivo non è stato possibile chiedere di poter far uso di un algoritmo che rappresenti lo stato dell’arte nelle analisi geospaziali.

Ciononostante, il modello di Optimal Interpolation è formalmente equivalente ad un altro modello di interpolazione noto con il nome di Kriging.

L’algoritmo di Kriging viene spesso utilizzato nei software di analisi geospaziele in quanto rappresenta un modo efficace per ottenere interpolazioni equivalenti al metodo di Optimal Interpolation. Essendo un metodo molto diffuso è anche ben documentato e consente lo sfruttamento di librerie software già esistenti. Per questo motivo si è deciso di usarlo come algoritmo di interpolazioni all’interno di questa analisi.

In particolare si è scelto di far uso di un Kriging bidimensionale, per le ragioni descritte in seguito.

* + - 1. IMPOSSIBILITY TO HAVE ALTIMETRIC DATA

Gli algoritmi di Kriging possono essere eseguiti in spazi dimensionali a due o tre variabili e in tempi computazionali accettabili in entrambi i casi.

L’ostacolo principale all’implementazione di un Kriging tridimensionale è dato però dalla natura del database fornito da AREU. I dati contemplano infatti la sola posizione geografica definita dalle coordinate X e Y ma mancano della componente altimetrica necessaria ad un’interpolazione tridimensionale.

* + - 1. LOW PRECISION OF EXTRACTED ALTITUDE

Per cercare di ovviare al problema di cui al punto 1) si è dunque cercato di procedere ad un’interpolazione del dato altimetrico facendo uso dell’API Open Elevation, un’alternativa open source concorrente al servizio Elevation offerto a pagamento da Google.

Per procedere all’estrazione del dato altimetrico si è innanzitutto scaricata la mappa SRTM250 (Shuttle Radar Topology Mission con precisione di 250m) generata dalla NASA.

La mappa consiste in un file formato .tiff che associa le coordinate nel formato (Longitudine, Latitudine) delle quali si vuole richiedere l’altitudine alle coordinate nel formato (X,Y) dell’immagine. Il colore di ogni pixel nell’immagine rappresenta il valore dell’altitudine nelle date coordinate.

Sfortunatamente la precisione a 250m risulta non sufficiente ad eseguire un’interpolazione corretta, soprattutto in aree ad altro gradiente altimetrico. Ad un’analisi comparativa, il modello utilizzante l’altezza interpolata come terza variabile di input risulta avere una precisione inferiore alla più semplice interpolazione bidimensionale. Per questa ragione si è scelto di utilizzare un modello di Kriging 2D.

Il risultato ottenuto è sicuramente da imputare alla mancanza di precisione nelle rilevazioni dell’altitudine ed è sicuramente risolvibile utilizzando mappe SRTM a risoluzione più elevata, che tuttavia non sono reperibili, se non sotto compenso, sul sito della NASA.

* + - 1. COMPUTATIONAL COMPLEXITY AND EXPLORATIVE ANALYSIS

Come detto, i modelli di Kriging bidimensionale e tridimensionale risultano pressochè equivalenti quando si tratta di operare su dati ricavati da una regressione altimetrica effettuata con una mappa la cui risoluzione non è sufficientemente fine. Ciò che cambia è invece la complessità computazionale dell’algoritmo di Kriging, che passa da O(n^2) a O(n^3), il che si traduce in un dilatarsi dei tempi d’esecuzione quantificabile in un 11X medio.

Essendo questa una tesi a carattere esplorativo non focalizzata sul metodo di interpolazione ma sulla rilevanza dei risultati, si è dunque deciso di optare per il modello più semplice al fine di rendere sostenibile l’interpolazione di un lag period di 5 giorni con rilevazioni ogni 12 ore per ognuno dei 9 fattori indagati.

* + 1. KRIGING
       1. SIMPLE KRIGING
       2. ORDINARY KRIGING
       3. UNIVERSAL KRIGING
    2. INTERPOLATION PROCEDURE
  1. INTERPOLATED DATA
     1. VARIOGRAM
     2. VARIOGRAM FITTING TO EXTRACT A MODEL
  2. VALIDATION
     1. VARIANCE AND PERCENTILES
     2. COMPARING THE REFERENCE MODEL

1. RESULTS
   1. FOUND TRENDS
   2. CORRELATIONS
2. RISK FUNCTION
3. CONCLUSION