

INFORMAZIONI PERSONALI

Valentina Minnetti

ESPERIENZA
PROFESSIONALE

(gennaio 2011 – in corso)

Ricercatore

Istat - Via A. Depretis 74, in Roma

(giugno 2020 – in corso)

Attività Varie

monitoraggio dei siti web degli istituti di statistica europei ed extra-europei riguardo aggiornamenti di strumenti statistici utili a misurare come la pandemia da COVID-19 impatta sulle statistiche economiche, sociali, etc. I risultati di un iniziale screening dei siti sono stati riportati nel lavoro: "Impacts of the COVID-19 on the statistical activities of the European National statistical institutes" (vedi pubblicazioni). Giudizio molto positivo del referente.

(dicembre 2019 – in corso)

Attività Varie

- 1. data mining: fase pre-processing

Analisi della qualità e capacità informativa di fonti dati, in particolare big data.

È usato software R.

- 2. data mining: fase post-processing

studio di nuovi metodi di analisi multi-criterio per la creazione di indicatori sintetici delle performance di modelli di classificazione binaria,

- 3. metodi di integrazione dati: uso di metodi probabilistici (in record linkage), metodi deterministici (query SQL), metodi di machine learning (supervisionato e non-supervisionato).

- Sono usati software SAS e R

(giugno 2019 – dicembre 2021)

Attività Varie

ESSNet - Big Data Pilots II – Work package Innovative Tourism Statistics (WPJ)

- Analisi dei dati con tecniche di data linking, dati web e da indagine.

- Sono usati i software SAS e Excel. Sostituire con le principali attività e responsabilità

(settembre 2017 - dicembre 2018)

Attività Varie

- Analisi di data quality dei dati sperimentali, ottenuti da sensori.

- È stato usato SAS

(dicembre 2016 - dicembre 2019)

Attività Varie

1. analisi di data mining applicati a dati sperimentali

2. analisi degli indicatori di performance dei modelli

3. studio di metodi di classificazione multi-classe per variabili economiche ATECO.

(aprile - novembre 2016)

Attività Varie

- studio di fattibilità dell'utilizzo del modello RDF a supporto della progettazione del registro attraverso la struttura linguistica delle ontologie.

Il parere positivo espresso sulla fattibilità ha contribuito al successo dello stesso.

(giugno 2015 - aprile 2016)

Attività Varie

- analisi di data quality della tabella QRCA (quality report card).

- analisi di data quality di fonti amministrative

- È stato usato Oracle/SQL plus.

(gennaio - aprile 2013)

Attività Varie

- analisi statistica multivariata degli indicatori di Comuni, Province, Regioni, ASL, Aziende Ospedaliere, relativi all'anno 2010, nel portale statistico della Pubblica Amministrazione.

- Sono stati usati SAS/BASIC, SAS/STAT, Excel.

(settembre - dicembre 2012)

Attività Varie

- creazione di data mart per alimentazione del Data Warehouse del 6° Censimento Generale dell'Agricoltura.

- È stato usato SAS/BASIC.

(febbraio 2011 - agosto 2012)

Attività Varie

- analisi dei dati da fonte amministrativa
- È stato usato SAS/BASIC, Oracle

(febbraio 2011)

Attività Varie

- definizione dei programmi dei corsi SAS 1 e SAS 2 svolti da giugno 2011 fino a tutto l'anno

(2005 al 2008)

Collaboratore informatico-statistico

Attività svolta: data entry.

- dal 2006 al 2008

Casa di cura S. Antonio da Padova S.R.L. – Via Mecenate 14, Roma

Casa di cura Città di Roma S.p.A. – Via Mairalchini 20, Roma

- dal 2006 al 2007

Casa di cura Samadi S.p.A. – Via di Grottarossa km 2.200, Roma

Aurelia'80 S.p.A. – Via Aurelia 860, Roma

- dal 2005 al 2007

European Hospital – Via Portuense 700, Roma

ISTRUZIONE E FORMAZIONE

(Aprile 2015)

Dottorato di Ricerca in Ricerca Operativa – ciclo XXV

Facoltà di Ingegneria dell'Informazione Informatica e Statistica, Sapienza Università di Roma

- Titolo: On the parameters of the Electre Tri method: a proposal for a new two phases procedure (vedi pubblicazioni)
- Supervisor: prof. Renato De Leone (Università di Camerino) Supervisor anno 2010: prof. Maurizio Vichi (Sapienza, Università di Roma)
- Parole chiave: metodi per il supporto alle decisioni multi-criterio - modelli di classificazione supervisionata - programmazione lineare.
- Sono stati usati i software SAS, R.
- Ogni anno ammessa al successivo con merito
- I risultati ottenuti, presentati alla conferenza AIRO 2011 sono stati molto apprezzati da esperti del settore, tanto da aver avuto la proposta di presentare il lavoro al meeting europeo MCDA'75, organizzato da Euro Working Group on Multicriteria Decision Aiding a Tarragona presso Universitat Rovira i Virgili (Catalogna).

(Luglio 2007)

Laurea specialistica in Statistica e Tecnologie Decisionali

Facoltà di Scienze Statistiche, Università degli Studi di Roma La Sapienza

- Titolo: Decisioni Multi-criterio e Analisi dei dati
- Votazione: 110/110 e Lode
- Relatore: prof. Paolo Dell'Omo ("Decisioni Multi-Criterio e Scelte Collettive")
- Riassunto: uso combinato di tecniche di data mining e un modello di decisione multi-criterio su dati di telefonia.
- Sono stati usati i software SAS/Enterprise Miner, SAS/BASIC e UTA-PLUS.

(Marzo 2005)

Master in Business Intelligence

Iresm Placido Martini in Via Machiavelli 70, Roma

Sono stati usati: SAS, MySQL, SAS macro

(Dicembre 2004)

Laurea in Statistica

Facoltà di Scienze Statistiche, Università degli studi di Roma La Sapienza

- Titolo: Nuovo strumento per il supporto alle decisioni: UTA-PLUS
- Votazione: 110/110 e Lode
- Relatore: prof. Paolo Dell'Omo ("Metodi e Modelli per il Supporto alle Decisioni")
- Riassunto: descrizione del modello di decisioni multi-criterio UTA e del corrispondente software UTA-PLUS.

(Dicembre 2002)

Attestato di frequenza del corso "Introduzione al SAS", tenuto dal Prof. D'Arcangelo ("Statistica per la ricerca Sperimentale")

Facoltà di Scienze Statistiche, Università degli studi di Roma La Sapienza

- (Settembre 2001) **Attestato di frequenza Corso inglese**
Proteo ED Services s.r.l. – Via Filippi 37, Roma
- (Settembre 2000) **Certificato di conoscenza della lingua inglese 1° livello**
Centro Linguistico di Ateneo, Università degli studi di Roma Tre, Roma
- (7 Luglio 2000) **Attestato di qualifica professionale di “Tecnico Automazione Uffici”, riconosciuto dalla Regione Lazio cofinanziato dal Fondo Sociale Europeo**
Centro di Formazione Professionale in Via Decio Azzolino 11, Roma
In aula 400 ore, stage 60 ore presso i Vigili del Fuoco, Via Genova 3/a Roma
- (22 Luglio 1996) **Diploma Maturità Scientifica**
Liceo Scientifico M. Malpighi, Roma.
Votazione: 52/60

COMPETENZE PERSONALI

Lingua madre Italiano

Altre lingue

	COMPRENSIONE		PARLATO		PRODUZIONE SCRITTA
	Ascolto	Lettura	Interazione	Produzione orale	
Inglese	B2	B2	B2	B2	B2

Competenze comunicative

Attitudini e competenze sociali:

- a) buona capacità di comunicazione/trasmettere efficacemente informazioni:
 - presentazioni a conferenze;
 - docenze
- b) buono spirito di gruppo:
 - seggio elettorale (numerose volte);
 - partecipazione a gruppi di lavoro

Competenze organizzative e gestionali

Attitudini e competenze organizzative:

- a) ottima capacità di coordinamento (presidente di seggio elettorale)
- b) buona capacità di organizzare eventi (sessioni di conferenze)
- c) buona capacità di gestione di piccoli gruppi di lavoro (scientifici)
- d) ottima capacità di gestione di progetti a distanza (tesi PhD)

Competenze professionali

Attitudini e competenze tecniche:

- a) ottima capacità di analisi di big data/milioni di dati
- b) buona capacità di creare nuove metodologie di analisi dati
- c) buona capacità di creazione di modelli ibridi provenienti da ambiti diversi

Competenze digitali

Attitudini e competenze informatiche:

sistemi operativi:

- a) ottima conoscenza di Windows
- b) conoscenza di base di Linux e Mac

programmi:

- c) ottima conoscenza dei programmi Office
- d) ottima conoscenza di programmi per il supporto alle decisioni multicriterio: UTA-Plus, Electre Tri,

Electre III/IV

e) buona conoscenza di SAS/(BASE, STAT, OR), SAS/Enterprise Miner

f) buona conoscenza di R

g) buona conoscenza di ITSM 2000

h) conoscenza di base dei programmi Weka, SPSS, MS Project

linguaggi di programmazione:

i) buona conoscenza di SQL in: Oracle SQL Plus, SAS, MySQL

j) conoscenza di base di PHP e HTML

k) buona conoscenza di LaTeX

ULTERIORI INFORMAZIONI

Altro

Attività di Consulenza

gennaio 2005

supporto all'elicitazione dei pesi dei criteri nel modello di classificazione Electre Tri, nell'analisi di supporto alle decisioni multi-criterio, utilizzato nella tesi PhD on Forestry Ecology "A Multi-Criteria Decision Framework : assessing synergies among the Rio conventions at forestry project level" di Rocío Dánica Córdor Golec (http://dspace.unitus.it/bitstream/2067/1977/1/rdcondor_tesid.pdf)

Attività di Docenza

marzo - maggio 2022

corso Data Quality and other issues of Official Statistics (6CFU), nel corso di Laurea Statistical Methods and Applications (Curriculum OFFICIAL STATISTICS) svolto su online, erogato dal Dipartimento di Scienze Statistiche, Sapienza Università di Roma

marzo - maggio 2012

corso "Laboratorio di Inferenza Statistica" (3 CFU), svolto presso il Dipartimento di Scienze Statistiche, Sapienza Università di Roma.
È stato usato R.

luglio 2010

corso "Matematica" (20h), svolto presso Liceo Scientifico Statale Torricelli, Roma

maggio - ottobre 2010

modulo "Laboratorio Statistico Informatico" (36h), nel corso I.F.T.S. "Tecnico Superiore dei Trasporti e dell'Intermodalità", svolto presso SEDIIN, Via Icilio 14, Roma.

Sono stati usati R e SAS/STAT.

maggio - ottobre 2010

modulo "Fonti statistiche per le economie locali" (12h), nel corso "Data Analyst per P.M.I.", svolto presso SEDIIN, Via Icilio 14, in Roma
Sono stati usati SAS/BASIC e Excel.

febbraio - luglio 2008

moduli:

"Organizzazione Aziendale e Gestione dei Processi" (26h);

"Performance logistiche: sistemi di misura, pianificazione e controllo" (50h);

"Modelli Stocastici per la Logistica Integrata" (11h)

nel corso I.F.T.S. "Tecnico Superiore nella Gestione Integrata della Catena Logistica",

svolti presso Accademia Informatica S.r.l., Viale Città d'Europa, 39, Roma.

Sono stati usati SAS/BASIC, SAS/OR e MS Project.

ottobre 2007 - aprile 2008

moduli:

“Statistica di Base” (6h);

“Strutture Dati e Algoritmi” (6h);

“Logistica” (10h);

“Laboratorio Statistico Informatico” (34h);

“SPSS” (24h).

nel corso I.F.T.S. “Tecnico Superiore della Logistica Integrata”, svolti presso il Liceo Statale Seneca, Via F. Albergotti, 35 in Roma Sono stati usati R, SPSS e SAS aprile (2005/2006/2007/2008/2009)

Lezioni su:

“Electre e AHP: casi di studio”;

“Introduzione al metodo UTA”.

“Metodo UTA e applicazione con UTA-Plus”

nel corso “Decisioni Multi-Criterio e Scelte Collettive” (prof. Paolo Dell’Olmo), presso la Facoltà di Scienze Statistiche, Università La Sapienza di Roma

ottobre 2005 - aprile 2006

ottobre 2004 - aprile 2005

ottobre 2003 - aprile 2004

moduli:

“Calcolo Combinatorio” (4h)

“Calcolo delle probabilità” (6h)

“Statistica” (6h in O5.A6) e (4h)

nel corso extra-scolastico “Laboratorio di Mathematica”, svolti presso Liceo Scientifico Statale Torricelli, Roma

Attività di Commissione

Giugno-settembre 2022

corso Data quality and other issues of Official Statistics (online)

giugno 2012

Membro di commissione d’esame del corso “Laboratorio di Inferenza Statistica”, presso il Dipartimento di Scienze Statistiche, Sapienza Università di Roma.

febbraio 2008

Membro di Commissione delle selezioni per il corso I.F.T.S. “Tecnico Superiore nella Gestione Integrata della Catena Logistica” presso Accademia Informatica s.r.l., Viale Città d’Europa, 39 in Roma

Aprile (2005/2006/2007/2008) - giugno 2009

Membro di Commissione d’esame del corso “Decisioni Multi-Criterio e Scelte Collettive”, con prof. P. Dell’Olmo, presso la Facoltà di Scienze Statistiche, Università di Roma La Sapienza

Pubblicazioni

paper su rivista/book/etc.

september 2020

Methods and analysis for combining web scraping data with data on tourist accommodations survey

(with M. Di Torrice, M. Amarone, L. Cavallo, M.T. Santoro, A. Simone)

In Proceedings of the Conference NTTS 2021, session Data-on-go, on march 2021.

Abstract: in this work, results of an initial experimental analysis is presented.

The web scraped data from tourism web sites are analyzed combined with the data of the administrative register of tourist accommodation establishments, for one region of Italy. Record linking methods are applied to the web data in order to estimate the

quality of them.

Maggio 2020

Impacts of the COVID-19 on the statistical activities of the European National statistical institutes

Abstract: In the observation window from 2020 april 23 and may 5, the activities (new statistical tools, variuos statistics, etc) of all the National statistical institutes (NSI), during the first wave of the pandemic are analyzed. Due to the difficulties of monitoring in real time the sites, a summarizing description of the statistical tools is presented, considering various areas (social, economic, etc) in order to put in evidence the different approaches of the NSIs to the numbers of the pandemic.

March 2017

On the UTA methods for solving the model selection problem

in "Optimization and Decision Science: Methodologies and Applications" – pp 59-67 – editors: A. Sforza and C. Sterle © Springer Proceeding in Mathematics & Statistics vol.217

Abstract: In this work, the multi-criteria decision analysis approach is proposed for solving the problem of model selection. Among the methods, the UTA methods realize a ranking of the models, from the best to the worst ones, by means of the comparisons of their global utility values. These values are computed by means of Linear Programming problems. The UTAs method is introduced as the improving method with respect to the UTA method. An example, that examines the performances of some classification models in the statistical experimental data is presented.

DOI 10.1007 / 978-3-319-67308-0 (indicizzato su Scopus, Google Scholar, etc)

September 2015

Electre Tri Machine Learning approach to the Record Linkage

(with R. De Leone)

in "Classification, (Big) Data Analysis and Statistical Learning", part of the Series Studies in Classification, Data Analysis and Knowledge Organization, pp 31-39 © Springer International Publishing AG2018

Editors: Mola F., Conversano C. and Vichi M.

Abstract: This work proposes, for the first time in the literature, the application of the method Electre Tri to solve the Record Linkage matching. The results of a preliminary experiment show that high accuracy can be achieved and more than 99% of the matches and nonmatches are correctly identified by using of the Electre Tri.

DOI 10.1007 / 978-3-319-55708-3 (indicizzato su Scopus, Google Scholar, etc)

April 2015

On the parameters of the Electre Tri method: a proposal of a new two phases procedure.

PhD thesis on Operational Research, Faculty of Information Engineering, Informatics and Statistics, Sapienza University of Rome, Italy.

Abstract: the aim of this thesis is to provide a new methodology for estimating the parameters of a classification model in the multi-criteria decision analysis context. In general, given a finite set of units and a finite set of classes, a classification model assigns each unit to one of the classes, a priori defined by the Decision Maker (DM). The new proposed methodology is based on an iterative process, in which the DM is supported in the worst case when the knowledge of the problem is not sufficient and no parameter is known.

The core of the methodology is the procedure in two phases. The first phase is the most important with the estimation of one parameter from which the other ones depend mathematically.

The innovation of this PhD thesis is the use of the linear programming (easy, that guarantees global solutions) instead of the nonlinear programming (hard, that guarantees local solutions) suggested by some authors in literature.

Moreover, this methodology is complete because all the parameters are considered, while in all the works in literature some parameters are not considered, because complicate the computations or are considered to be fixed.

The application contexts are in various sectors: military, environmental, medical, geographic, energetic, etc, and for solving other mathematical problems.

The application presented in the thesis is the use of the multi-criteria classification model to solve the record linkage (RL) problem, for the first time in literature.

The term RL means the problem of recognizing identical persons, objects, events in different data sets, but also the recognition of the duplicates records in the same data set.

In this thesis, only the results of a first experiment are inserted in order to test the validation of the multi-criteria classification model for solving the RL problem.

Software used: SAS/OR, SAS/BASE, R, MPL.

February 2015

A new distance measure for solving Record Linkage with Electre Tri.

Technical report, Department of Statistical Science in Sapienza University of Rome.

Abstract: In the Record Linkage, the aim is to find the same person in two different presentations a conference/meeting/

files or is to individuate duplicates in the same file. In order to compare two units of the Cartesian Product space, a distance measure must be used. In this work, a new distance measure, is formulated, that is the generalization of an existing measure.

September 2013

The Estimation of the Parameters in Multi-Criteria Classification Problem: The Case of the Electre Tri Method

(with R. De Leone)

in "Analysis and Modelling of Complex Data in Behavioral and Social Sciences", part of the Series Studies in Classification, Data Analysis and Knowledge Organization, pp 93-101 © Springer International Publishing Switzerland 2014.

Editors: Vicari D., Okada A., Ragozini G. and Weihs C..

Abstract: the aim of this work is the estimation of the parameters of the multi-criteria classification model Electre Tri by means of the linear programming.

A methodology is suggested for estimating all the parameters, considering the absence of any information on the parameters.

DOI 10.1007/978-3-319-06692-9 (indicizzato su Scopus, Google Scholar, etc)

October 2011

Il data warehouse del 6° Censimento generale dell'agricoltura

(con M. Dimitri, F. Consentino, M. Passacantilli)

in "Gli Analytics come motore per i Big Data la Ricerca e il Sistema Paese" a cura di A. Di Ciaccio, W. Lanzani. ARACNE Editrice S.r.l. di Roma – I edizione, settembre 2013

Sintesi: in questo poster sono elencate con breve descrizione le procedure SAS usate per la creazione di data mart che hanno alimentato il data warehouse del VI censimento generale dell'agricoltura.

ISBN 978-88-548-6324-8

Maggio 2011

Uso applicativo degli archivi amministrativi

Sintesi: in questo lavoro viene mostrato come una base dati corretta e completa abbia una capacità informativa notevole, attraverso un esempio applicativo dalla base dati anagrafica dell'anagrafe tributaria.

presentazioni a conferenze/meeting/

Settembre 2017 – Sorrento (NA), Italy

'On the UTA methods for solving the model selection problem'
In Session: "Data Science 2" of the 47th ODS - AIRO Conference

ottobre 2015 – S. Margherita di Pula (CA), Italy
'Electre Tri-Machine Learning Approach to the Record Linkage problem'
In Contributed Session: "Data Mining" of the 10th CLADAG Meeting

settembre 2015 – Pisa, Italy
'Electre Tri-Machine Learning Approach to the Record Linkage'
In Session: "Classification & Forecasting" of the 45th AIRO conference

luglio 2013 – Roma, Italy
'An application of Calibration Sampling problem: the case of the ElectreTri method'
In Technical Session "Methods of Multi-Criteria Decision Analysis " of the 26th EURO – Informs conference

settembre 2012 - Vietri sul Mare (SA), Italy 'The estimation of the parameters of the Electre Tri method: a proposal of a new two step procedure'
In Session: "Classification and Clustering Problems" of the 43rd AIRO conference

settembre 2012 – Capri (NA), Italy
'The Classification Problem in Multi-Criteria Decision Analysis: the Electre Tri method'
In Contributed Session "Advances in Classification and Clustering 2" of the JCSCLADAG conference

luglio 2012 – Vilnius, Lithuania
'On the estimation of the parameters of Electre Tri model in multi criteria ordinal sorting problem: a proposal of new approach in two phases'
In Technical Session "Preference Learning 2" of the 25th EURO conference

settembre 2011 – Brescia, Italy
'New Approach to Estimate the Parameters of Electre Tri Model in the Ordinal Sorting Problem'
In Session: "Regression and Multiple Criteria Decision Aid" of the 42nd AIRO conference

giugno 2011 – Camerino, Italy
'Support Vector Machines and Linear Programming Based Techniques for Classification and Regression'
In Lecture of the 1st Scientific Day of the School of Science and Technology

settembre 2010 – Roma, Italy
'Approccio multi-criterio al problema di classificazione'
Nel Meeting "Comunicazione e ricerca: i giovani si incontrano", presso Dipartimento di Scienze Statistiche, La Sapienza Università di Roma

ALLEGATI

Sostituire con la lista di documenti allegati al CV. Esempi:

- copie delle lauree e qualifiche conseguite
- attestazione del datore di lavoro

Dati personali

Autorizzo il trattamento dei miei dati personali ai sensi del Decreto Legislativo 30 giugno 2003, n. 196 "Codice in materia di protezione dei dati personali".

I sottoscritt_ dichiara di essere consapevole che il presente *curriculum vitae* sarà pubblicato sul sito istituzionale dell'Ateneo, nella Sezione "Amministrazione

trasparente", nelle modalità e per la durata prevista dal d.lgs. n. 33/2013, art. 15.

Data

f.to