

Corso di Dottorato in Economia e Statistica XXXV ciclo, a.a. 2019/2020

n. 1 Borsa finanziata su fondi “Dipartimenti di Eccellenza” e vincolata al progetto “Studi teorici e applicati sul funzionamento dei sistemi economici dei paesi Europei” (curriculum Economia)

Ai candidati si richiede la capacità di proporre e sviluppare ricerche nell’ambito dell’analisi teorica e applicata dell’economia dell’Unione Europea. A questo proposito, si segnala che presso il DEMS sono attualmente attive le seguenti linee di ricerca, a cui i candidati possono fare riferimento: 1) Revisione della governance Europea per accrescere efficacia ed efficienza dell’interazione tra governi nazionali e organi sovranazionali. 2) Analisi econometrico/statistica del processo di convergenza tra i paesi e le regioni europee. Integrazione nominale, reale e finanziaria. Interazioni macro-finanziarie e cicli boom-bust. 3) Definizione di politiche macroeconomiche volta a una maggiore condivisione dei rischi attraverso: a) un incremento del bilancio dell’unione (fondo europeo di disoccupazione); b) il coordinamento delle politiche fiscali nazionali. 4) Monitoraggio di rischiosità del sistema finanziario dell’unione e identificazione di linee guida comuni in materia di assunzione del rischio dei vari attori nei sistemi finanziari. 5) Politiche di integrazione dei mercati e politiche industriali. 6) Analisi delle politiche commerciali e migratorie dell’Unione Europea. Resta ovviamente aperta la possibilità per i candidati di proporre nuove linee di ricerca.

n. 3 Posti di dottorato con Percorso Executive, curriculum Statistica

n.1 posto riservato a dipendenti UniCredit S.p.A.

Titolo/Argomento della Tesi

Holistic approach to Operational Risk: Issues, Solutions, and Decision Making

Profilo

Il candidato svolgerà un percorso di studio e di ricerca volto a completare le competenze statistiche con moderne metodologie di statistical learning, machine learning, web analytics e text mining. Tali conoscenze verranno man mano integrate a modelli statistici più classici per costruire una metodologia rigorosa in grado di fornire allarmi precoci di possibili perdite finanziarie.

Progetto

Operational Risk losses usually are classified into standardized categories, such as Basel Event Types and Business Lines, however unstructured descriptions are often available, containing valuable information. Analyzing the text data associated to Operational Risk losses, using text mining techniques, can provide a more granular classification, and can facilitate the discovery of new

phenomena, that would otherwise remain hidden in the multitude of events, sources and types of data. Additional Operational Risk related data can be retrieved in the web and processed using the web analytics techniques. Moreover, applying machine learning algorithms (e.g. neural networks) to the above mentioned data can enable deep-dives on specific topics, can provide the OpRisk management function with a high-level view of millions of events, and can raise early warnings on related data, that could generate losses.

Ore di formazione esterna (in Università)

Il percorso di studio del candidato prevede la partecipazione a corsi di statistical learning, machine learning, web analytics e text mining. Per i primi due anni di dottorato si concorderanno almeno due corsi da seguire, così come si prevede la partecipazione a summer school e seminari per gli anni successivi, scelti in base al formularsi sempre più preciso degli interessi di ricerca. Si prevede che il candidato possa recarsi all'estero per periodi di studio.

Esami obbligatori

Non se ne prevede alcuno a priori; in base all'offerta di corsi disponibili nel dottorato EcoStat si indicherà al candidato di svolgere gli esami più pertinenti al proprio percorso formativo.

n.1 posto riservato a dipendenti RAMDAC S.R.L.

Capacità informativa dei profili di navigazione dei cookies.

Utilizzare i profili di navigazione dei cookies di editori e agenzie media per valutare la capacità di questi profili di prevedere i comportamenti e identificare meglio i soggetti. Tutto questo in ambiente big data con milioni di cookies e con strumenti di profilazione di avanguardia come classificatory e motore semantico e con strumenti di machine learning in ambiente Hadoop+Spark.

n.1 posto riservato a dipendenti DATA REPLY vincolata al progetto di ricerca

Titolo/Argomento della Tesi

I principali argomenti sui quali si fonderà la tesi sono:

- Classificazione con dati ad alta dimensionalità
- Elaborazione di dati ad alta dimensionalità

Profilo

Data scientist with advanced statistical skills

Progetto

Il progetto si fonda sulla ricerca e lo sviluppo di modelli di classificazione e previsione basandosi su dati ad alta dimensionalità come le immagini, utilizzando sia gli approcci della statistica classica sia le più moderne tecniche di machine learning e deep learning. Una componente fondamentale sarà la ricerca di metriche per valutare la bontà di adattamento di tali modelli.

Ore di formazione esterna

50% di lavoro, 50% Università (primo anno di PHD)

Esami obbligatori

- Probability
- Stochastic process
- Statistical Inference
- Statistical Learning
- Data Management
- Computational Statistics

n.1 scholarship reserved to employees of CGNAL S.r.l. related to research project

News sentiment and financial markets: A framework for developing sentiment based trading strategies using machine learning algorithms.

Profile

Data Scientist, CGnal Srl

The project idea consists in analyzing unstructured data from news feeds and social networks to create different sentiment signals related to companies, commodities, market indexes and/or financial indicators in general. These signals will be used afterwards to model the relationship between markets and sentiment with the objective of creating trading strategies. The first part considers two steps: * First, the analysis of the textual sources to create a model that generates sentiment scores for each document. Possible alternatives for this task are: - Dictionary-based algorithms, where a sentiment lexicon is created with words and their respective score, then this is used as an input in a model to calculate a final document score. - Transfer learning based, using word vectorization algorithms trained on huge amounts of textual data as base layers to develop a deep learning architecture to calculate sentiment for each document. In this task is important to consider the source and the context. For example, a regular sentiment lexicon not created specifically for financial terms would not work as well, some words generally used in financial circles

like "bull" or "bear" have specific meanings that the developed model needs to understand correctly. On the other hand, sources like Twitter, have peculiarities that more "traditional" textual sources don't have. The length of the message is limited, there's a prevalence of non-word structures, like emojis and also the presence of communication conventions specific to the platform like hashtags, mentions and cashtags that need to be processed and recognized in specific ways by the developed algorithm. * Second, the creation of sentiment signals and features based on the aggregation of these scores. It's necessary to recognize which documents to use for creating each sentiment signal, define a relevant time interval and then aggregate them in a meaningful way for the modeling step. The second part, is modeling of the sentiment signal against asset returns for obtaining long or short signals. This requires a feature engineering and selection for each asset analyzed and then a traditional modeling step. Possible approaches for this:

- Bayesian structural time series approaches - Traditional machine learning approaches, based on supervised learning: different ensemble approaches (bagging, boosting).
- Neural Network approaches like, recurrent neural networks or convolutional neural networks (traditional or dilated for sequences).
- Reinforcement learning approaches that learn from trades how to interpret in a better way the sentiment signals.