**PROPOSTE DI TIROCINIO**

**1. ORCHESTRAZIONE DI *CONTAINER* IN AMBIENTE *PRIVATE CLOUD***

L’iniziativa ha come obiettivo lo studio delle prassi e delle tecnologie per l’esercizio in ambiente *enterprise* di strumenti e sistemi per l’orchestrazione di microservizi in *private cloud,* attraverso l’impiego delle piattaforme *OpenStack e kubernetes.*

Attraverso lo studio e l’utilizzo di quanto realizzato in Banca d’Italia, il tirocinante avrà anzitutto modo di apprendere le competenze di base relative ai concetti del *Software Defined Data Center* (SDDC), al paradigma *cloud* e alle tecnologie connesse ai *container*; successivamente le sue attività si concentreranno sulla comparazione di diverse modalità di utilizzo di *kubernetes* su ambiente *OpenStack,* così da produrre come obiettivo finale un processo automatico di configurazione e gestione che ne garantisca la resilienza.

Il tirocinante collaborerà con il *team* incaricato di seguire gli sviluppi e la gestione dell’SDDC in Banca d’Italia, alternando il lavoro individuale con quello di gruppo.

**Competenze acquisite al termine del tirocinio:**

Una o più delle seguenti:

* conoscenza dell’ambiente operativo Linux e dei suoi principali strumenti;
* conoscenza delle tecnologie impiegate in ambito *cloud* (*OpenStack* e *kubernetes*);
* competenze sullo sviluppo e gestione di servizi IT in ambienti basati su macchine virtuali e *container*;
* sviluppo di capacità di lavoro in gruppo e in autonomia.

**Requisiti richiesti** (oltre a quelli già specificati)**:**

* laurea magistrale in Ingegneria informatica, Ingegneria delle tecnologie della comunicazione e dell’informazione, Matematica, Fisica, Scienze computazionali.

**Conoscenze richieste** (titolo preferenziale)**:**

* conoscenza del sistema operativo Linux;
* nozioni di infrastrutture di *server*, rete e *storage*;
* conoscenza dei sistemi di virtualizzazione (*virtual machine* e *container*) in ambiente Linux;
* capacità di utilizzare risorse in ambiente *cloud*;
* conoscenza del *cloud* *management* *system* *OpenStack*;
* conoscenza del *software* *Kubernetes*.

**2. TECNOLOGIA “*R”* PER L’ELABORAZIONE DI DATI STATISTICI**

L’iniziativa ha l’obiettivo di sviluppare un modello per il supporto alle attività di realizzazione della Rilevazione economica e tecnologica sull’IT nel sistema bancario, condotta annualmentedalla CIPA - Convenzione Interbancaria per l’Automazione (www.cipa.it). Le attività riguarderanno l’innovazione degli strumenti tecnologici attualmente a supporto del processo di elaborazione/analisi dei dati e produzione della reportistica, con particolare riferimento all’adozione di modelli e metodi *open* (quali l’ambiente R); potranno inoltre essere studiati ulteriori metodi di rilevazione e/o scenari di analisi.

Il tirocinante collaborerà con il *team* che produce le indagini statistiche, approfondendo l’attuale processo di raccolta e analisi dei dati nonché i metodi adottati per la produzione dei grafici e della reportistica. La sua collaborazione avrà principalmente l’obiettivo di migliorare il processo in essere mediante l’introduzione di strumenti e tecnologie innovative. In particolare, il tirocinante contribuirà alle attività, già avviate dalla Divisione, di migrazione nell’ambiente di programmazione R delle elaborazioni finalizzate all’analisi dei dati e alla produzione dei grafici. Le tecnologie utilizzate in tale ambito sono infatti per la maggior parte di tipo tradizionale e proprietario, basate su strumenti consolidati quali SAS e Excel.

**Competenze acquisite al termine del tirocinio:**

Una o più delle seguenti:

* conoscenza di un processo tipico di raccolta e analisi dei dati di tipo statistico sull’ambito IT nel sistema bancario;
* competenze sulla realizzazione di un progetto di migrazione tecnologica, in ambito statistico, da soluzioni proprietarie a soluzioni *open source*;
* competenze sul linguaggio di analisi dei dati R e sulla programmazione a oggetti.

**Requisiti richiesti** (oltre a quelli già specificati)**:**

* laurea magistrale in Ingegneria informatica, Ingegneria delle tecnologie della comunicazione e dell’informazione, Matematica, Fisica, Scienze computazionali.

**Conoscenze richieste** (titolo preferenziale)**:**

* conoscenza dei *database* relazionali (es. Oracle) e degli strumenti di *office automation* per la reportistica (es. Excel);
* nozioni di programmazione in un linguaggio di alto livello (es. Python, PHP, Java);
* conoscenza dei linguaggi R, SAS o altri per l’analisi statistica dei dati;
* conoscenza di ambienti e librerie per l’analisi dei dati e la produzione di reportistica (preferibilmente SAS e R) nonché conoscenza di librerie (es. D3, Gplot2, Plotly, etc.) per la produzione di grafici.

**3. *CHATBOT* PER LA FRUIZIONE DI SERVIZI AZIENDALI**

L’illusione di interagire con una macchina come se fosse un essere umano è un problema che, seppur formalizzato dal matematico Alan Turing negli anni ‘50, è sempre più di interesse sia per la comunità scientifica sia per aziende di ogni tipo. Colossi come Google, Amazon, Samsung e Microsoft sono attivamente impegnati per fornire soluzioni che facilitano l’interazione uomo-macchina al fine di migliorare la fruizione di ogni genere di servizio.

In questo contesto, *chatbot* e *Virtual Personal Assistant* si comportano come dei veri e propri assistenti personali, permettendo di acquistare voli, prenotare il ristorante più vicino, ottenere le previsioni del tempo, gestire gli *smart devices* di casa o ricevere assistenza tecnica. Le aziende stanno incominciando ad utilizzare soluzioni di questo tipo per agevolare le interazioni con i clienti e con i propri dipendenti, a volte anche in sostituzione dell’operatore umano.

Il tirocinante dovrà affrontare lo studio delle principali tecnologie alla base dei *chatbot*, con particolare riferimento agli approcci basati su regole e su reti neurali di apprendimento. L’attività prevede la messa in atto delle tecnologie apprese al fine di progettare e sviluppare un *chatbot* aziendale che permetta di ottimizzare la fruizione di informazioni/servizi accessibili mediante la rete *Intranet* aziendale; in particolare, i sistemi di *car pooling*, *car sharing*, *bike sharing*, mensa, navette per il trasporto, organigramma interno e gestione dell’orario di lavoro.

**Competenze acquisite al termine del tirocinio:** Una o più delle seguenti:

* conoscenza di algoritmi di *machine learning* e *deep learning*;
* modellazione architetturale di soluzioni *software*/infrastrutturali;
* capacità di programmare in Python.

**Requisiti richiesti** (oltre a quelli già specificati)**:**

* laurea magistrale in Ingegneria informatica, Ingegneria delle tecnologie della comunicazione e dell’informazione, Matematica, Fisica, Scienze computazionali.

**Competenze richieste** (titolo preferenziale)**:**

* predisposizione allo sviluppo *software* e fluidità con almeno un linguaggio di programmazione;
* predisposizione allo studio di pubblicazioni scientifiche.

**4. ANALISI DI SISTEMI DINAMICI COMPLESSI**

L’approccio tradizionale allo studio dei fenomeni economici come sistemi complessi è stato a lungo quello riduzionista (*equation-based modeling*): modellazione di una versione semplificata del sistema al fine di renderlo analiticamente trattabile. Questa strategia ha permesso di conseguire risultati teorici rilevanti. Un approccio alternativo è la *simulation-based modeling*, che costituisce un’area attiva di ricerca e consente di modellare sistemi complessi con un più elevato grado di realismo e, in prospettiva, di produrre previsioni ed analisi di *policy* più affidabili.

Nell’ambito delle attività della Divisione Ricerca sulle Tecnologie Avanzate della Banca d’Italia, è stato individuato un insieme di modelli di mercati reali e finanziari presenti in letteratura, che sono poi stati studiati approfonditamente ed implementati. Sono inoltre in fase di sviluppo versioni evolutive di questi modelli.

Sotto la guida di un ricercatore-tutor della Banca d’Italia e nell’ambito di un team di progetto, il tirocinante avrà la possibilità di studiare modelli realistici di mercati reali e finanziari. Avrà inoltre la possibilità di contribuire all’implementazione dei modelli, utilizzando i più importanti *framework* di riferimento (es. *NetLogo*, *Java Agent Development Framework*). Si potrà far ricorso a simulazione su larga scala con l’adozione di paradigmi e strumenti scalabili di programmazione concorrente (es. *actor-based programmin*g e *Akka*).

Il tirocinante parteciperà inoltre allo studio, alla sintesi e alla presentazione dei risultati delle simulazioni mediante l’utilizzo dei principali pacchetti statistici.

**Competenze acquisite al termine del tirocinio:**

Una o più delle seguenti:

* progettazione e implementazione di modelli multi-agente;
* conoscenza di *framework* per l’implementazione di simulazioni di sistemi ad agenti;
* conoscenza di paradigmi e strumenti di programmazione concorrente;
* capacità di analisi e sintesi di dati multivariati con pacchetti statistici (R, Stata, Matlab, etc.);
* conoscenza del funzionamento di mercati reali e finanziari;
* modellazione realistica di sistemi economici.

**Requisiti richiesti** (oltre a quelli già specificati)**:**

* laurea magistrale in Ingegneria informatica, Ingegneria delle tecnologie della comunicazione e dell’informazione, Matematica, Fisica, Scienze computazionali, Scienze economiche.

**Competenze richieste** (titolo preferenziale)**:**

* predisposizione allo sviluppo e fluidità con almeno un linguaggio di programmazione tra: Java, C#, Scala, Python;
* conoscenze di base di statistica descrittiva ed inferenziale;
* esperienza nell’utilizzo di un pacchetto statistico;
* interesse nel funzionamento dei sistemi economici, conoscenze di base di teoria economica.

**5. SVILUPPO DI SISTEMI DI *REASONING* AUTOMATICO PER *BIG DATA* E *KNOWLEDGE GRAPHS***

Sempre più aziende moderne e “*Internet giants*” (ad es. Facebook, Amazon, Google) desiderano modellare e mantenere la conoscenza sotto forma di *corporate knowledge graphs* e gestirla attraverso *knowledge graph management systems* (KGMS), sistemi in grado di rispondere a interrogazioni e svolgere ragionamenti complessi in modo automatico.

Allo stesso tempo, un KGMS deve garantire alte prestazioni e scalabilità in ambito Big Data, con una complessità computazionale accettabile, ed essere in grado di interfacciarsi con una molteplicità di sistemi: basi di dati relazionali e NoSQL *stores*, librerie di *machine learning* e *analytics*, sorgenti *web*. La Divisione Ricerca sulle Tecnologie Avanzate del Dipartimento IT della Banca d’Italia è attiva in questo ambito e contribuisce allo sviluppo di un importante KGMS, in stretto contatto con l’Università di Oxford e aziende di riferimento nel settore.

Il tirocinio ha come obiettivo la progettazione e lo sviluppo di *knowledge graphs* per il supporto di importanti applicazioni del dominio finanziario.

Esso prevede lo studio e l’applicazione di tecnologie e metodologie di *machine learning* e inferenza logica su *knowledge graphs* in ambito Big Data. Le attività saranno collocate nell’ambito di un progetto aziendale strutturato; il lavoro sarà svolto in *team* secondo pratiche agili e sotto la costante guida di ricercatori-tutor della Banca d’Italia.

Il tirocinante svilupperà soluzioni alternativamente in Python o in Java e avrà l’opportunità di utilizzare molteplici sistemi e librerie per il *data management* tra cui: DBMS relazionali, Neo4J, Dgraph, NetworkX, Cytoscape, Spark MLib, GraphX, Spark Streaming, ecc.

Inoltre, il tirocinante potrà partecipare alla scrittura di articoli scientifici, con la possibilità di collaborare con esperti internazionali del settore e presentare il proprio lavoro in prestigiose sedi scientifiche e del mondo finanziario.

**Competenze acquisite al termine del tirocinio:**

* utilizzo e sviluppo di tecnologie di *reasoning* automatico per Big Data;
* utilizzo e sviluppo di tecniche di *machine* *learning* e di inferenza automatica.

**Requisiti richiesti** (oltre a quelli già specificati)**:**

* laurea magistrale in Ingegneria informatica, Ingegneria delle tecnologie della comunicazione e dell’informazione, Matematica, Fisica, Scienze computazionali.

**Competenze richieste** (titolo preferenziale)**:**

* basi di dati relazionali;
* competenze di programmazione (Java/Python).

**6. SPERIMENTAZIONE DI TECNICHE DI *MACHINE LEARNING* SU DATI STRUTTURATI E NON STRUTTURATI**

La Banca d’Italia tratta in modo automatico una grande quantità di dati, provenienti dal mondo economico e finanziario, con molteplici finalità. Le sfide e le opportunità poste dal moderno contesto Big Data, rendono opportuno il ricorso a tecniche innovative che supportino i processi caratterizzati da *human in the loop*. In particolare, è di rilievo fornire supporto agli analisti nelle fasi di processo ripetitive identificando automaticamente quei casi critici che necessitano l’analisi dettagliata di un esperto.

In quest’ambito, il tirocinante contribuirà alla realizzazione di un sistema per il supporto all’operatività degli analisti della Banca; in particolare, svilupperà modelli di classificazione automatica, *clustering* e *mining* basati su *machine learning* e *deep learning*.

Sarà centrale la realizzazione di modelli in grado di acquisire congiuntamente la conoscenza da sorgenti di dati siano esse strutturate (tabelle) che non strutturate (testo libero in linguaggio naturale).

Il tirocinio sarà collocato nell’ambito di un progetto aziendale strutturato; il lavoro sarà svolto in *team* secondo pratiche agili e sotto la costante guida di ricercatori-tutor della Banca d’Italia.

Il contesto tecnologico di riferimento è: Python, scikit-learn, gensim, MongoDB, Tensorflow.

**Competenze acquisite al termine del tirocinio:**

* conoscenza delle principali tecniche di *text mining*;
* conoscenza di modelli e strumenti di *machine learning* e *deep learning*.

**Requisiti richiesti** (oltre a quelli già specificati)**:**

* laurea magistrale in Ingegneria informatica, Ingegneria delle tecnologie della comunicazione e dell’informazione, Matematica, Fisica, Scienze computazionali.

**Competenze richieste** (titolo preferenziale)**:**

* predisposizione allo sviluppo di codice prototipale e fluidità con il linguaggio Python;
* fondamenti di *machine learning* o *data mining*.

**7. TECNICHE DI *WEB SCRAPING* PER L’ACQUISIZIONE DI DATI SOCIETARI**

La Banca d’Italia gestisce numerose basi di dati contenenti informazioni relative alle società italiane. Tali informazioni sono utilizzate con varie finalità, dallo svolgimento dei compiti istituzionali alla conduzione di analisi a carattere economico.

Le tecniche di analisi più avanzate (ad esempio, quelle basate su *Knowledge Graph*) beneficiano fortemente di processi di *data augmentation*, che possano completare e – in alcuni casi – correggere la base di conoscenza. Quasi sempre queste informazioni vengono estratte dal *web* o dai *social network*.

In questo ambito, il tirocinante collaborerà alla progettazione e alla realizzazione di un sistema che, a partire da alcuni dati strutturati relativi ad una certa società, reperisca automaticamente dal *web*, tramite tecniche di *web-scraping*, alcune *feature* di interesse aggiuntive, integrando ed eventualmente correggendo i dati già a disposizione. Il sistema dovrà essere in grado di estrarre dati strutturati da milioni di siti *web* in maniera automatica.

Il tirocinio sarà collocato nell’ambito di un progetto aziendale strutturato; il lavoro sarà svolto in *team* secondo pratiche agili e sotto la costante guida di ricercatori-tutor della Banca d’Italia. Inoltre, il tirocinante potrà partecipare alla scrittura di articoli scientifici, con la possibilità di collaborare con esperti del settore e presentare il proprio lavoro in prestigiose sedi scientifiche e del mondo finanziario.

**Competenze acquisite al termine del tirocinio:**

* conoscenza di tecniche avanzate *web scraping* e *data extraction*.

**Requisiti richiesti (**oltre a quelli già specificati)**:**

* laurea magistrale in Ingegneria informatica, Ingegneria delle tecnologie della comunicazione e dell’informazione, Matematica, Fisica, Scienze computazionali.

**Competenze richieste** (titolo preferenziale)**:**

* predisposizione allo sviluppo e fluidità con i linguaggi Java o Python.