

PROGETTO FORMATIVO: Attività didattica programmata – Insegnamenti previsti

I ANNO		
N	Insegnamento	Descrizione
1	<i>Data science e metodi di intelligenza artificiale</i> Ore: 10 Verifica Finale: SI	Il corso si pone l’obiettivo di studiare il complesso processo della Data Science sull’analisi e interpretazione dei dati grazie agli strumenti del machine learning. Un progetto di creazione di modelli di dati prevede le seguenti fasi: 1. Pianificazione; 2. Costruzione di un modello di dati; 3. Valutazione del modello; 4. Spiegazione del modello; 5. Monitoraggio del modello. La data science, infatti, combina più campi tra loro, tra cui in particolare la statistica e l’intelligenza artificiale, per analizzare i dati raccolti dal Web, dagli smartphone, dai clienti, dai sensori e da altre fonti per ricavare insight utili. Le aziende oggi basano le proprie attività su una preziosa raccolta di dati, spesso conservati in database e data lake, per lo più intatti. I dati raccolti e archiviati offrono vantaggi in termini di trasformazione ad aziende in tutto il mondo, solo se si è in grado di interpretarli. Infatti, l’elevata dimensionalità dei dati introduce sfide computazionali e statistiche uniche: le aziende non si accontentano più di estrarre informazioni dettagliate dai loro archivi ma richiedono l’applicazione di modelli predittivi complessi quali gli analytics. In particolare, nel corso si approfondirà il concetto di machine learning, i cui algoritmi trasformano le informazioni in conoscenza fruibile. Questo fatto rende il machine learning particolarmente adatto all’attuale epoca dei big data. 1. Definizione di Big Data; 2. Fonti diverse e Statistica: confronto tra scienza basata sui dati (data driven) e basata sui modelli (model based); 3. Classificazione dei Big Data: human generated data, process mediated data, machine generated data; 4. Fonti offline dei Big Data e IoT: beacon, sensori biometrici, digital signage; 5. Introduzione agli algoritmi di machine learning, quali ad esempio apprendimento pigro e probabilistico, clustering, alberi decisionali, algoritmi di previsione, Metodi di black-box: reti neurali e macchine a vettori di supporto.
2	<i>Statistica e probabilità</i> Ore: 10 Verifica Finale: SI	Il corso si propone di familiarizzare i dottorandi con le metodologie basilari della ricerca scientifica, portandoli gradualmente alla scoperta delle problematiche connesse alla costruzione delle informazioni statistiche. Obiettivo del corso è di fornire ai dottorandi le competenze per analizzare in modo autonomo i dati empirici, padroneggiando l’uso dei principali test statistici, sia di base che avanzati. Saranno introdotte, inoltre, nozioni avanzate del calcolo delle probabilità, dell’analisi dei dati e dell’inferenza statistica. Al termine del corso il dottorando sarà in grado di applicare queste conoscenze a problemi scientifici e applicazioni in cui è necessario modellizzare fenomeni aleatori e condizioni di incertezza. In particolare, il ruolo del calcolo delle probabilità consiste nel misurare, principalmente, sulla base dei risultati relativi al campione, l’attendibilità di circostanze o eventi non osservati. La raccolta, l’elaborazione e l’interpretazione dei dati statistici conducono, in generale, ad uno stato di informazione in cui la valutazione quantitativa dell’incertezza, mediante il concetto di probabilità, è diversa rispetto alla situazione iniziale. I dottorandi apprendono come applicare il calcolo delle probabilità e le tecniche inferenziali a problemi reali e come derivare risultati teorici in modo formale. Viene acquisita consapevolezza sui vantaggi e sui limiti delle metodologie trattate, in modo da saper valutare criticamente quali siano gli strumenti più adatti nelle specifiche situazioni da sottoporre ad analisi. Il corso sarà incentrato sui seguenti argomenti, non esclusivi: 1. Logica e classificazione dei modelli statistici; 2. Modelli lineari generalizzati; 3. Inferenza per i modelli lineari generalizzati; 4. Processi stocastici.
3	<i>Programmazione per data analytics</i>	La data science è una moderna disciplina che combina aree della matematica e informatica per estrarre conoscenza e valore dai dati. La conoscenza ricavata dai dati è alla base di previsioni e decisioni aziendali e governative, di scoperte scientifiche, di studi di fenomeni sociali, ma anche di applicazioni di intelligenza artificiale che, dalla computer vision alla elaborazione del linguaggio naturale, stanno rivoluzionando l’economia e la società. Obiettivo del corso è insegnare a trasformare database

DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE” – XLII ciclo

Curriculum: Big data management per la transizione digitale

I ANNO		
	Ore: 10 Verifica Finale: SI	disordinati e dati non strutturati in dataset omogenei e ordinati attraverso l'impiego degli strumenti essenziali del linguaggio Python e R per il data mining, il machine learning, la network analysis e l'elaborazione del linguaggio naturale. Il dottorando imparerà ad acquisire dati numerici e testuali eterogenei da fonti diverse come semplici file locali, database e il Web e a strutturarli nel modo più utile, a trasformarli, a visualizzarli e a modellarli, utilizzando le funzionalità delle principali librerie per il machine learning. Le analisi dei Big Data presuppongono tuttavia una piattaforma software che consente di distribuire complessi compiti di computing su una grande quantità di nodi. Tra le tecnologie in grado di soddisfare le esigenze dei Big Data, c'è Apache Hadoop, sviluppato in ambiente open source, considerato uno dei fattori abilitanti la diffusione dei Big Data Analytics. Inoltre, per l'integrazione e la gestione dei dati, il classico archivio informatico di dati strutturati raccolti dai sistemi operazionali aziendali e da fonti esterne, è insufficiente. Oggi il Data Warehouse è integrato con il Data Lake, un ambiente di archiviazione dei dati nel loro formato nativo. Agli strumenti di storage, si affiancano tipologie innovative di gestione dei dati. Non solo database relazionali, ma anche basi di dati che superano o rinnovano il linguaggio SQL (si parla per questo di database NoSQL o NewSQL), al fine di migliorarne le performance.
4	<i>Modelli statistici avanzati e data management</i> Ore: 15 Verifica Finale: SI	Il corso fornisce le competenze e le tecniche necessarie per sfruttare i dati, per conoscere le criticità e per supportare in modo adeguato il processo decisionale. L'obiettivo è formare dottori di ricerca, con robuste competenze sulle metodologie e sui modelli di analisi dei dati, che nelle singole funzioni aziendali saranno coinvolti nella valutazione analitica di informazioni quantitative e qualitative relative a dati aziendali. Non solo gestione di dati e informazioni, caratteristiche proprie del Data Scientist, ma anche capacità di sapersi relazionare con le diverse aree del management per una efficace comunicazione delle analisi svolte e dei risultati ottenuti al fine di valutare e supportare le decisioni aziendali. Il dottore di ricerca deve saper prendere decisioni basate su evidenze empiriche dando risposte appropriate alle richieste informative aziendali e saper estrarre le informazioni rilevanti dai dati di grandi dimensioni (big data). I recenti sviluppi tecnologici e dei software open source hanno aumentato la potenza computazionale, favorendo la fruibilità di algoritmi di elaborazione dati più efficaci, generando un aumento della richiesta di soluzioni predittive più avanzate e sofisticate attraverso l'utilizzo di algoritmi di Machine Learning. Le tecniche di Machine Learning appartengono ad una famiglia di algoritmi che racchiudono insieme la statistica applicata e le scienze informatiche. Il dottore di ricerca dovrà essere in grado di saper riorganizzare dati multidimensionali con struttura complessa e analizzarli utilizzando metodologie di statistica multivariata, quali ad esempio: 1. Integrazione e riduzione dei dati; 2. algoritmi di classificazione e regressione; 3. algoritmi supervisionati e non supervisionati, 4. modelli di sopravvivenza; 5. analisi fattoriale mediante SEM; 6. serie storiche.
5	<i>Data Science e Scienze Regionali</i> Ore: 15 Verifica Finale: SI	I recenti progressi tecnologici stanno contribuendo alla produzione di una raccolta di dati incredibilmente ampia e in rapida accelerazione, che va sotto il nome Big Data. Nell'era dei big data tutte le nostre interazioni e transazioni quotidiane generano dati. Questi dati ci permettono di esaminare i fenomeni urbani e regionali con tecniche di analisi che prima non erano possibili. Tuttavia, nonostante l'enorme potenziale dei Big Data per le scienze regionali, il loro uso e la loro applicazione in questo contesto presentano problemi e criticità di non poco conto. Nel corso saranno altresì discusse le sfide e le prospettive future per l'uso dei big data nelle scienze regionali e nei sistemi di governance del territorio. Quando si parla di disponibilità e pubblicità dei dati, infatti, della loro accessibilità e riusabilità, si parla di trasparenza amministrativa e di Open Data. Se la Pubblica Amministrazione è supportata da dati e informazioni aggiornate, può analizzare al meglio le situazioni e intervenire in modo mirato in settori quali la mobilità, l'ambiente e la qualità della vita dei cittadini. Lo sviluppo urbano passa oggi dalla governance dei dati. Un approccio organico e integrato ai dati grazie alle tecnologie abilitanti dell'IoT, dell'intelligenza artificiale e del Cloud, consente l'interoperabilità, la condivisione dei dati in tempo reale e il governo dei big data. Nel corso saranno affrontati i seguenti argomenti, non esclusivi: I parte 1. Economia del territorio: dimensione spaziale e territoriale; 2. Analisi esplorativa e strutture dei dati spaziali; 3. Review della letteratura sulle principali linee di ricerca della statistica spaziale

I ANNO		
		II parte 4. Misure di autocorrelazione ed eterogeneità spaziale; 5. Modelli di regressione spaziale avanzati. III parte 6. Geo Data Analytics per le Decisioni Aziendali; 7. Costruzione e visualizzazione di carte geografiche per dati spaziali: la georeferenziazione.

DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE” – XLII ciclo

Curriculum: Big data management per la transizione digitale

II ANNO		
N	Insegnamento	Descrizione
1	<i>Intelligenza artificiale per data-driven economy</i> Ore: 15 Verifica Finale: SI	La Data Driven Economy significa un’economia generata da un nuovo tipo conoscenza, fondata sui Big Data ovvero sulla capacità di utilizzare al meglio la quantità di informazioni che, ogni giorno, viene prodotta, basandosi sull’Intelligenza Artificiale. È una Algorithmic Economy, basata su dati statistici, fondato su un approccio più Agile, capace di attingere anche a dati nuovi, in cloud o sui social, ad esempio, per applicare l’intelligenza degli algoritmi e costruire servizi a misura di utente. Il corso si pone l’obiettivo di offrire una presentazione completa dei metodi statistici e di apprendimento automatico utili per analizzare e prevedere serie di dati di grandi dimensioni e dinamicamente dipendenti. Saranno discussi algoritmi di clustering e di classificazione delle serie storiche e i metodi di riduzione delle dimensioni in presenza di dipendenza dinamica (modelli fattoriali dinamici). Il corso affronterà anche argomenti modelli di previsione in presenza di Big Data, utilizzando algoritmi di machine learning, quali: 1. Reti neurali e Deep learning; 2. Foreste casuali; Boosting, Bagging 4. Nowcasting. 5. AI: interpretazione vs spiegabilità 6. Time series, big data e AI
2	<i>Tecniche di analisi dei big data e di reti sociali</i> Ore: 15 Verifica Finale: NO	L’ecosistema tecnologico legato ai Big Data è estremamente vario, tanti sono i nuovi strumenti e numerose le nuove sfide da affrontare. Un’architettura dei Big Data, tecnologicamente all’avanguardia è strutturata sulle diverse fasi di gestione del dato: dall’impostazione della Data Strategy al Data Management passando per le fasi di Data Integration e Data Visualization. Il corso si propone di fornire agli studenti conoscenze teoriche e competenze applicative nell’ambito delle moderne metodologie di analisi dei dati e dell’intelligenza artificiale. In particolare, il percorso formativo mira a sviluppare la capacità di comprendere, modellare e analizzare fenomeni complessi attraverso strumenti statistici, algoritmi di apprendimento automatico e tecniche avanzate di data science. Saranno approfonditi i principi e le applicazioni del linguaggio naturale e dell’intelligenza artificiale, con particolare attenzione ai modelli utilizzati per l’elaborazione automatica dei testi e delle informazioni non strutturate. Il corso introdurrà i fondamenti della data science e dello statistical learning, fornendo gli strumenti necessari per l’analisi esplorativa, la costruzione di modelli predittivi e la valutazione delle loro prestazioni. Particolare attenzione sarà dedicata all’analisi funzionale dei dati, finalizzata allo studio di fenomeni rappresentati da curve, segnali e serie temporali continue. Infine, saranno affrontate le metodologie di network graph analysis per l’analisi delle relazioni tra individui, organizzazioni o entità complesse, consentendo di interpretare strutture reticolari e dinamiche di interazione in diversi contesti applicativi. Al termine del corso gli studenti saranno in grado di utilizzare strumenti avanzati per l’analisi dei dati, applicare tecniche di intelligenza artificiale e machine learning a problemi reali e interpretare criticamente i risultati ottenuti a supporto dei processi decisionali.
3	<i>Nuove frontiere dei big data nella statistica spaziale</i> Ore: 15 Verifica Finale: NO	I metodi spaziali rappresentano solo una parte dei metodi quantitativi; tuttavia, nell’ultimo decennio, nella scienza dei dati si è assistito ad un crescente interesse per i sistemi informativi geografici e per i metodi spaziali quantitativi. La quantità crescente di informazioni raccolte è di fatto costituita da dati spaziali, contenenti informazioni di geolocalizzazione. Si tratta, ad esempio, dei dati sul percorso di ritorno dal lavoro o dalla passeggiata o della posizione dei negozi o dei punti di servizio più visitati. Ma anche le informazioni pubblicate su Twitter, le query inserite nella casella di ricerca, le voci o le foto pubblicate sui social network contengono molto spesso dati di geolocalizzazione. Da qui la comparsa del concetto di big data spaziale. I Big Data possono rappresentare, in tal senso, una nuova opportunità per sviluppare analisi più accurate e geo-referenziate, accanto a quelle più tradizionali. Obiettivo del corso è di fornire informazioni sui concetti base e avanzati propria della statistica spaziale. In particolare, il corso prevede: 1. Modelli di analisi dei processi di punto spaziali

II ANNO		
		2. Processi spazio-tempo 3. Applicazione dei machine learning ai dati spaziali: algoritmi supervisionati e non supervisionati 4. Geomarketing, Geocomputation e Spatial Data Science 5. Metodi e modelli di analisi spaziale avanzata per la valutazione dell'impatto territoriale delle decisioni pubbliche. La seconda parte del corso introduce il concetto di Economic complexity. La disciplina Economic Fitness and Complexity rappresenta una delle metodologie usate per la valutazione dell'efficacia e dell'impatto dei piani di sviluppo dei paesi europei. Tale approccio tiene conto della competitività di una nazione misurando il livello di diversificazione e di complessità dei prodotti esportati da quel paese e si concretizza in un'analisi innovativa ed interdisciplinare basata su dati e metodi scientifici moderni dell'area dei Sistemi Complessi.

DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE” – XLII ciclo

Curriculum: Big data management per la transizione digitale

III ANNO		
N	Insegnamento	Descrizione
1	<i>Opportunità e sfide tecnologiche per un mondo interconnesso</i> Ore: 10 Verifica Finale: NO	La crescita esponenziale di dispositivi connessi sempre più miniaturizzati e con sempre maggiore capacità computazionali e la nuova generazione in grado di supportare un'elevata densità di dispositivi e garantire prestazioni avanzate in termini di latenza e affidabilità, stanno creando un mondo iperconnesso, fatto di oggetti di vario tipo (sensori, droni, auto a guida autonoma) favorendo casi d'uso innovativi, opportunità e sfide nella nostra società e nella vita quotidiana. Il successo dell'avvento di questo scenario iperconnesso dipende dalla capacità di integrare tecnologie come Internet of Things, reti avanzate per la comunicazione Machine-to-Machine, veicoli/droni connessi, cloud computing, edge computing, raccolta/ disseminazione dati e social network. Questo corso presenta una panoramica delle tecnologie abilitanti di questi nuovi scenari e il ruolo giocato e le opportunità portate dai Big Data e Intelligenza Artificiale. Inoltre, il corso discuterà vari casi di studio in questo campo al fine di presentare le sfide aperte e comprenderne le potenzialità di ricerca (interdisciplinare). Particolare attenzione sarà dedicata all'applicazione dell'Intelligenza Artificiale nell'Industria 5.0 e nei modelli sostenibili di produzione e consumo, evidenziando il ruolo delle tecnologie digitali nell'ottimizzazione dei processi industriali, nell'efficienza energetica e nella gestione intelligente delle risorse. In particolare, il corso sarà incentrato sui seguenti argomenti: 1. Reti ad alta velocità servizi avanzati: integrazione IoT, cloud e reti di nuova generazione, architetture di rete e servizi; 2. Reti industriali e veicolari: integrazione con reti 5G/6G ed edge computing; 3. Raccolta e diffusione di dati pervasivi, ruolo dell'intelligenza artificiale; 4. Discussione e presentazione di casi di studio (E-health, Smart cities, Industry 5.0).
2	<i>Intelligenza artificiale sostenibile</i> Ore: 10 Verifica Finale: NO	Le enormi quantità di dati generate negli ultimi vent'anni e lo sviluppo impetuoso della capacità di calcolo, uniti attraverso l'Intelligenza Artificiale (IA), permette di analizzare, capire e migliorare la gran parte delle attività umane e delle aziende. In particolare, in campo economico l'IA consente alle aziende di ottimizzare l'uso delle risorse riducendo al minimo gli sprechi e nel riciclare persino le rimanenze della produzione. D'altra parte, lo sviluppo e l'utilizzo di IA è spesso associato anche ad un possibile impatto negativo sulle persone e sul pianeta per i possibili utilizzi che se ne possono fare, soprattutto in assenza di un chiaro quadro normativo che ne definisca i limiti e i principi etici da rispettare. Obiettivo del corso è di evidenziare come l'IA possa effettivamente costituire una leva virtuosa per il perseguimento della sostenibilità sociale, economica ed ambientale. In particolare, nel corso saranno mostrate come le soluzioni di IA possano avere un impatto significativo sulla conservazione delle risorse, la riduzione delle emissioni, la gestione dei flussi di traffico e dei relativi rischi, il rafforzamento dell'economia circolare e la prevenzione dei disastri naturali. Nel corso verrà anche evidenziato anche l'aspetto di sostenibilità della IA. Infatti, le soluzioni di IA possono richiedere infatti una grande quantità di energia elettrica per il loro funzionamento, che potrebbero incentivare l'immissione di CO2 nell'atmosfera. Se l'energia, però, deriva da fonti rinnovabili, l'impatto in termini di emissioni si riduce a zero: in quest'ultimo caso, l'IA può rappresentare un incentivo alla transizione verde del settore energetico. Ciò è ben descritto nella Strategia nazionale per l'Intelligenza Artificiale dove si evidenzia come l'IA sarà un alleato fondamentale per accelerare la transizione ecologica, un pilastro del piano di ripresa e resilienza dell'Italia e degli sforzi di ripresa dell'Unione Europea.

DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE” - XLII ciclo

Curriculum: Big data management per la transizione digitale

PROGETTO FORMATIVO: Altre attività didattiche (seminari, attività di laboratorio e di ricerca, formazione interdisciplinare, multidisciplinare e transdisciplinare)

N	Insegnamento	Descrizione
1	<i>Seminari</i>	I ANNO: PRESENTAZIONE AREA DI RICERCA DEL DOTTORATO - Questo seminario è pensato per introdurre gli studenti alla tematica di ricerca oggetto del corso al fine di renderli pienamente consapevoli dell'ambito di studio nel quale sono chiamati a muoversi e che presenta anche carattere di interdisciplinarietà. Inizialmente, verrà presentata la tematica Big Data e Intelligenza Artificiale, con le molteplici sfaccettature, i trend in atto e le principali sfide da affrontare. Inoltre, saranno discussi i due curricula del dottorato con lo scenario di riferimento, le sfide specifiche e gli ambiti applicativi di riferimento. Verranno altresì presentati i professori del collegio dei docenti e il programma degli insegnamenti sia per la formazione specifica sui temi inerenti la tematica del corso sia per la formazione trasversale. Parimenti, saranno presentate le aziende coinvolte nel dottorato, presso le quali i dottorandi svolgeranno parte della loro formazione, ivi compresi i progetti di ricerca. Infine, saranno fornite informazioni di carattere logistico sulla piattaforma didattica dell'Ateneo.
2	<i>Valorizzazione e disseminazione dei risultati, della proprietà intellettuale e dell'accesso aperto ai dati e ai prodotti della ricerca</i>	I ANNO: VALORIZZARE E DISSEMINARE I RISULTATI DELLA RICERCA - Questo corso è dedicato alla gestione della ricerca sotto l'aspetto di valorizzazione disseminazione dei risultati della ricerca. L'obiettivo principale del corso è fornire agli studenti un metodo per diffondere e presentare in modo efficace il lavoro svolto e i risultati ottenuti. Verranno presentati e discussi con gli studenti degli esempi di presentazioni orali e articoli scientifici nel campo della Ingegneria, Statistica ed Economia evidenziandone i tratti comuni tra gli ambiti scientifici e caratteristiche peculiari nonché differenze legate alla modalità di disseminazione tra presentazione a conferenza o articolo su rivista. Verranno svolte esercitazioni pratiche e lavori di gruppo attraverso i quali i partecipanti svilupperanno le competenze necessarie per redigere e tenere presentazioni orali e scrivere articoli scientifici. Il programma si articola in: • Preparare e tenere una presentazione orale: preparazione del discorso, dalla stesura alla presentazione orale, struttura, stile e tempi della presentazione orale, le regole d'oro del parlare in pubblico, diversificazione della presentazione orale a seconda del destinatario; prova pratica seguendo le regole da “Coimbra three Minutes Competition rules”; • Praticare scrittura scientifica: le regole d'oro di redazione di un articolo scientifico e organizzazione in un gruppo di lavoro; strumenti di editing anche cooperativi; esempi di buona e cattiva scrittura scientifica; prova pratica di scrittura di un articolo scientifico; • La struttura e lo stile della scrittura scientifica: linee guida delle riviste; principali norme editoriali; processo di revisione fra pari; come gestire le richieste di Editori e Revisori; esempio di lettera di risposta ai revisori.
3	<i>Gestione della ricerca e della conoscenza dei sistemi di ricerca europei e internazionali</i>	I ANNO: COSTRUZIONE DI UNA PROPOSTA PROGETTUALE - L'obiettivo principale del corso è fornire una conoscenza del sistema della ricerca italiana, europea e internazionale con i soggetti coinvolti (pubblici e privati) e relativi programmi di ricerca e finanziamento. Inoltre Verranno forniti elementi utili per la valorizzazione dei risultati della ricerca nella partecipazione a bandi competitivi nei vari programmi di ricerca e in linea con gli elevati standard di valutazione. Inoltre verranno quindi fornite linee guide per interpretare correttamente il bando e individuarne i fattori chiave, valutare la corrispondenza dell'idea con l'argomento del bando, evidenziare i punti di forza dell'idea progettuale, soddisfare i criteri presenti nel bando. Riguardo alla generazione di impatto, verranno considerati non solo quelli di natura tecnologica, ma anche di natura economica, industriale, sociale e ambientale. Il programma si articola in: • Sistema e programmi di ricerca italiani, europei e internazionali: principali soggetti finanziatori, schemi di sovvenzione tra programmi quadro e bandi competitivi (PRIN, COST, Horizon Europe, ERC) e programmi di scambi internazionali; strumenti per la promozione dell'uguaglianza di genere e attinenza ai principi di etica della ricerca; analisi del processo di valutazione; • Costruire una proposta di ricerca: valutazione dell'idea progettuale rispetto alla tematica e requisiti del bando; analisi del grado di innovatività rispetto allo stato dell'arte; analisi dell'impatto; costruzione del partenariato: profili e ruoli previsti; • Scrivere una proposta di successo: costituire un consorzio funzionale ai criteri del bando; presentare l'idea progettuale, diagrammi di Pert e di Gantt per costruire un piano di lavoro efficace; costruzione del budget; strategie per impostare una corretta gestione della proprietà intellettuale e dell'accesso aperto ai dati nella proposta; esempi concreti di buone e cattive pratiche.
4	<i>Perfezionamento informatico</i>	I ANNO: STRUMENTI INFORMATICI PER LA RICERCA - Obiettivo del corso è presentare l'utilità di alcuni strumenti informatici utili per la ricerca dei dottorandi, al fine di presentare i loro risultati. Nel seminario sarà spiegato come acquisire dimestichezza con le basi dati comunemente utilizzate nella ricerca, nonché con i

DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE” - XLII ciclo

Curriculum: Big data management per la transizione digitale

N	Insegnamento	Descrizione
		principali strumenti di analisi degli stessi. Saranno presentate le modalità di gestione e organizzazione delle differenti tipologie di database, formule e fogli di lavoro, importazione e gestione di dati con diversi formati, elaborazione di statistiche descrittive, tabelle avanzate, pivot, grafici. Per realizzare grafici di funzioni, oltre ad Excel, saranno presentati all'interno del seminario alcuni programmi ad hoc quali Gnuplot o MAFAPlotter, tool gratuiti e open source disponibile per tutti i principali sistemi operativi, che consentono di visualizzare e tracciare grafici in 2D e in 3D partendo da funzioni matematiche. Altro strumento utile è Draw, appartenente alla suite Apache OpenOffice, che permette di realizzare diagrammi di flusso. Draw permette di esportare i grafici in formato pdf, jpeg, png, xml e html; il programma consente inoltre di salvare i lavori direttamente nei cloud (Dropbox, Drive e OneDrive). Verranno altresì presentati modalità e protocolli per il trasferimento dati, quali ad esempio ftp. A livello di scrittura, per la redazione di documenti, oltre al programma word, sarà descritto il linguaggio LaTeX, apprezzato per scrivere, ad esempio, articoli scientifici o documenti contenenti formule matematiche. Infine, per la presentazione dei propri risultati, nel seminario si tratterà delle funzionalità avanzate di PowerPoint per la creazione di presentazioni professionali. In particolare, verrà spiegato come gestire testo, immagini, loghi, oggetti complessi, colori e grafici per comunicare dati, analisi, piani di progetto, ecc.
5	<i>Perfezionamento linguistico</i>	I, II, III ANNO: INGLESE SCIENTIFICO - Il corso intende fornire elementi aggiuntivi di uso della lingua inglese in ambito scientifico e accademico. Il corso prevede due diversi parti. La prima è finalizzata all'acquisizione della consapevolezza di strutture linguistiche fondamentali per la scrittura e la pubblicazione di articoli scientifici in lingua inglese, nonché delle abilità linguistiche che sono necessarie durante la partecipazione interattiva in contesti accademici internazionali (convegni, seminari, ecc.). La seconda parte intende fornire competenze avanzate di lingua per una interazione efficace e per una maggiore adattabilità e flessibilità in contesti organizzativi complessi che caratterizzano sia la comunità accademica che il mondo del lavoro. In particolare, verranno forniti strumenti linguistici avanzati per la negoziazione, lavoro cooperativo in team, espressione di concetti complessi, e comunicazione interpersonale costruttiva ed efficace.
6	<i>Gestione della ricerca e della conoscenza dei sistemi di ricerca europei e internazionali</i>	II ANNO: PROMUOVERE E MONITORARE L'IMPATTO DELLA RICERCA - Questo seminario è dedicato alla gestione della ricerca sotto l'aspetto di promozione e monitoraggio dell'impatto della ricerca svolta nell'era digitale. Durante il seminario verranno dati elementi per orientarsi nell'ecosistema degli editori scientifici di collane e riviste e inoltre verranno presentati i maggiori database citazionali e i principali indicatori bibliometrici e identificativi univoci, offrendo un'analisi degli aggregatori e social network accademici, per comprendere come valutare e promuovere la propria produzione e attività scientifica. Inoltre verranno dati elementi di valutazione su come scegliere le destinazioni delle pubblicazioni e di caratterizzazione delle riviste (reputazione e impact factor, riviste predatorie).
7	<i>Seminari</i>	II ANNO: INFRASTRUTTURE DELLA RICERCA - Il seminario ha l'obiettivo di presentare le infrastrutture di ricerca come strumento fondamentale per realizzare sperimentazioni che richiedono grandi quantità di dati e di risorse, ad esempio, computazionali. Rientrano tra le infrastrutture di ricerca sia grandi attrezzature scientifiche che risorse quali collezioni, archivi, banche dati, ma anche infrastrutture virtuali (e-infrastructures) di dati, di calcolo, di comunicazione. Verranno mostrate le principali infrastrutture per la sperimentazione e la ricerca, in particolare quelle che offrono grandi capacità computazionali e facilità di programmazione specie per machine learning. Inoltre, verrà presentata la e-infrastruttura di ricerca europea per la sperimentazione su larga-scala che coinvolge più tipi di risorse quali cloud, reti wireless e 5G, IoT, big data, banchi di server. Oltre alla vasta gamma di risorse e tipologie di testbed, verranno presentati anche gli strumenti per l'accesso remoto alle attrezzature e risorse distribuite, nonché i servizi messi a disposizione per la sperimentazione remota, come ad esempio sonde di monitoraggio, strumenti di programmazione, collezione e memorizzazione di statistiche.
8	<i>Valorizzazione e disseminazione dei risultati, della proprietà intellettuale e dell'accesso aperto ai dati e ai prodotti della ricerca</i>	III ANNO: FARE SCIENZA APERTA OGGI: DALL'OPEN ACCESS ALL'OPEN DATA - L'obiettivo del seminario è fornire linee guide per la gestione e l'accesso aperto ai dati/prodotti della ricerca e al tempo stesso nelle questioni relative alla protezione della proprietà intellettuale e gli aspetti legali dell'editoria ad accesso aperto. Verrà introdotto il concetto di Open Data il concetto di dati FAIR (Findable, Accessible, Interoperable, Reusable) come prerequisito essenziale per la pubblicazione e la condivisione dei propri dati. Verrà trattato il concetto di Open Access e verranno presentati i recenti sviluppi nel panorama delle pubblicazioni ad accesso aperto nonché le sfide che i giovani ricercatori devono affrontare in questo ambito. Il programma si articola in:

DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE” - XLII ciclo

Curriculum: Big data management per la transizione digitale

N	Insegnamento	Descrizione
		● Open Data e Proprietà intellettuale: linee guida su Open Science nei programmi di ricerca europei ed internazionali, come utilizzare i repository di dati, aspetti di condivisione e pubblicazione dei dati della ricerca; aspetti legali per la gestione dei dati personali e la protezione della proprietà intellettuale nell'editoria ad accesso aperto, come il diritto d'autore, le licenze, modalità di riutilizzo dei dati e materiali esistenti; ● Gestione aperta dei dati e dati FAIR: ciclo di vita dei dati (dalla documentazione e immagazzinamento, all'utilizzo fino alla loro pubblicazione, protezione e riuso), come scrivere un piano di gestione dei dati, come rendere FAIR i propri dati seguendo gli standard esistenti e come valutare la FAIRness dei set di dati esistenti; ● Open Access: caratteristiche e vantaggi dell'accesso aperto ai prodotti della ricerca, linee guida su Open Access nei programmi di ricerca europei ed internazionali come trovare riviste open access per il proprio lavoro e differenza rispetto a riviste su piattaforme tradizionali, licenze per la pubblicazione di dati e software, implicazioni dell'accesso a pubblicazioni tramite piattaforme di pubblicazione alternative.
9	<i>Valorizzazione e disseminazione dei risultati, della proprietà intellettuale e dell'accesso aperto ai dati e ai prodotti della ricerca</i>	III ANNO: PROTEZIONE E VALORIZZAZIONE DEI RISULTATI DELLA RICERCA: TUTELE E STRUMENTI - L'obiettivo del corso è introdurre gli studenti nelle questioni relative alla tutela delle proprie idee e ai principali strumenti per la protezione dei risultati della ricerca e alla loro valorizzazione. Il corso intende rendere i giovani ricercatori consapevoli di cosa sia la proprietà intellettuale, come si genera e come si tutela. Si approfondiranno in particolare le caratteristiche delle invenzioni e dei brevetti. Inoltre, verranno illustrate le norme che regolano la titolarità dei risultati, il regolamento d'Ateneo in materia di proprietà intellettuale e industriale e i servizi offerti ai ricercatori per la tutela dei risultati innovativi frutto della loro ricerca. Il corso si propone inoltre di fornire le prime indicazioni di quali siano le possibilità per la valorizzazione della proprietà intellettuale e dei risultati della ricerca sul mercato. Si affronteranno le seguenti tematiche: ● Protezione della proprietà intellettuale: varie forme di proprietà intellettuale, proteggersi con il diritto d'autore: normativa e ambito di applicazione; diritto d'autore e diritto industriale: oggetto e differenze fondamentali; Struttura dei brevetti e fondamenti di ricerca brevettuale; ● Protezione del software: normativa vigente, modalità di protezione delle invenzioni implementate tramite computer; requisiti di validità, differenze ed vantaggi e svantaggi della protezione mediante brevetto o diritto d'autore; ● Valorizzazione risultati della ricerca sul mercato: valorizzazione tramite brevetto e/o costituzione di startup innovative, relazioni tra ricerca e aziende per il trasferimento tecnologico dei risultati della ricerca, regolamento in materia di Universitas Mercatorum.
10	<i>Principi fondamentali di etica, uguaglianza di genere e integrità</i>	III ANNO: ETICA E INTEGRITA' DELLA RICERCA - Questo seminario ha l'obiettivo di presentare i principali aspetti di etica nonché delle implicazioni di natura etica della ricerca e presentati gli obiettivi e le regole di funzionamento delle istituzioni che si occupano di ricerca scientifica in fatto di etica. Il seminario, inoltre, si propone di introdurre il concetto di integrità nella ricerca cioè l'insieme dei principi e dei valori etici, dei doveri deontologici e degli standard professionali sui quali si fonda una condotta responsabile e corretta da parte di chi svolge (finanzia) o valuta la ricerca scientifica. L'applicazione dei principi e dei valori e il rispetto della deontologia e degli standard professionali sono garanzia della qualità stessa della ricerca e contribuiscono ad accrescere la reputazione e l'immagine pubblica della scienza, con importanti ricadute sulla comunità scientifica e sulla società. Verranno presentati i principi e le responsabilità richieste ad ogni ricercatore durante tutto il processo della ricerca, dalla pianificazione alla pubblicazione, fornendo consigli pratici su come affrontare questioni complesse. Il seminario esplorerà gli standard a cui tutti i ricercatori devono tendere in modo efficiente e coinvolgente. Si affronteranno le seguenti tematiche: ● Etica nella ricerca: comitato etico: ruolo e funzioni; quando sottoporre una proposta di progetto al comitato etico; etica nei progetti che coinvolgono esseri umani; considerazioni di base sull'etica dei dati della ricerca e di protezione dei dati quando si lavora con i dati di ricerca che ha coinvolto esseri umani; ● Integrità della ricerca: principi dell'integrità della ricerca; conseguenze di una violazione di questi principi; come creare consapevolezza sull'integrità della ricerca e sulle responsabilità dei ricercatori nel condurre l'attività di studio; principali strumenti e standard di riferimento.
11	<i>Seminari</i>	III ANNO: REALIZZARE LA TRANSIZIONE DIGITALE VERDE - Il seminario analizzerà il tema dello sviluppo sostenibile inteso come sviluppo che soddisfa i bisogni del presente senza compromettere la possibilità delle generazioni future di soddisfare i propri bisogni. E' uno sviluppo che deve rispettare la cosiddetta regola delle tre E: 1. Ecologia: con ecologia si fa riferimento alla capacità dell'ambiente di mantenere le sue funzioni nel corso del tempo. Le principali funzioni dell'ambiente sarebbero quella di fornire risorse (dirette e indirette) e smaltire rifiuti;

DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE” – XLII ciclo

Curriculum: Big data management per la transizione digitale

N	Insegnamento	Descrizione
		2. Equità: il concetto di equità a cui si fa riferimento è quello di equità intergenerazionale, nel senso che le generazioni diverse godono del medesimo diritto di usufruire delle risorse del pianeta; 3. Economia: in questo caso ci si riferisce all’economia come alla capacità di creare una crescita duratura degli indicatori economici, generare reddito e lavoro e sostenere nel tempo le popolazioni, valorizzarne le specificità territoriali e allocarne efficacemente le risorse. Altro focus del seminario sarà costituito dal tema della transizione ecologica, ovvero dell’attitudine al risparmio energetico e alla sostenibilità ambientale che si conferma in Italia come un requisito di base per entrare nel mercato del lavoro. Secondo i dati di Unioncamere, nel 2021 le imprese hanno richiesto al 76,3% delle assunzioni programmate – pari a oltre 3,5 milioni di posizioni – competenze green, e nel 37,9% dei casi con un grado di importanza per la professione elevato. Il seminario, infatti, approfondirà anche le competenze dei green jobs, secondo i dati de Il Sistema Informativo Excelsior. I Green jobs sono caratterizzati, rispetto alle altre professioni, da una richiesta sempre più intensa di problem solving, di capacità di gestire soluzioni innovative e di elevate competenze matematiche e informatiche, oltre che digitali.
12	<i>Principi fondamentali di etica, uguaglianza di genere e integrità</i>	III ANNO: UGUAGLIANZA DI GENERE NELLA RICERCA E NELLA IMPRENDITORIA - Il seminario affronta il tema dell’uguaglianza di genere nel settore della ricerca e dell’innovazione che rappresenta un obiettivo primario della Commissione Europea e un obbligo giuridico previsto dal programma quadro dell’UE (Reg 1291/2013), con concrete conseguenze anche per il sistema accademico italiano. Il programma Horizon 2020 ha promosso in maniera significativa l’uguaglianza di genere, sostenendo l’equilibrio nei gruppi di ricerca, nei panel di supervisione e finanziando progetti specifici di sostegno alla politica di uguaglianza di genere. Nel seminario saranno illustrati alcuni dati che maggiormente danno prontezza della disuguaglianza di genere nell’università italiana e che fanno riferimento alla distribuzione dei docenti tra le diverse figure professionali accademiche. Tra il 2008 e il 2018, benché i professori ordinari siano, per tutte le aree scientifiche, a prevalenza maschili, possiamo cogliere un generale trend di crescita della presenza femminile dal 2008 al 2018. Gli incrementi maggiori si sono riscontrati nelle aree STEM. In particolare, Ingegneria industriale e dell’informazione ha registrato quasi un raddoppio. Tuttavia, il presupposto di partenza è la persistenza di un significativo gender gap, relativo alla presenza femminile nella ricerca e nei percorsi di formazione e carriera di interesse scientifico: il seminario partendo dal Rapporto She figures 2021 (Publications Office of the EU (europa.eu) evidenzierà eventuali percorsi di disparità di genere a partire dal conseguimento del titolo di dottorato di ricerca. In particolare, l’attenzione sarà incentrata sulla correlazione tra donne in possesso di un titolo di dottorato di ricerca e aumento di donne ricercatrici, soprattutto nelle aree di interesse scientifico.
13	<i>Valorizzazione e disseminazione dei risultati, della proprietà intellettuale e dell’accesso aperto ai dati e ai prodotti della ricerca</i>	III ANNO: COMUNICARE LA RICERCA NELL’ERA DEI SOCIAL MEDIA E PUBLIC ENGAGEMENT - In questo seminario verranno affrontati i vari aspetti legati alla comunicazione della ricerca rivolta a un pubblico generalista, e a fornire elementi per presentare in modo efficace la ricerca nell’era dei social media in un mondo della comunicazione sempre più affollato e competitivo. Infatti, saper produrre conoscenze scientifiche o innovazioni è cosa diversa da saperle comunicare in modo efficace. Sempre più spesso però bisogna saper spiegare il proprio lavoro a finanziatori, politici, imprenditori, giornalisti o cittadini, e convincerli. Dopo un’introduzione sul public engagement e sulla terza missione della ricerca (definizione, tipologie ed esempi, valutazione), il seminario fornirà un excursus sulla divulgazione scientifica in Italia e le principali tendenze di divulgazione scientifica anche su social media. Verrà fornita una panoramica dei mezzi di comunicazione e degli interlocutori non appartenenti alla comunità scientifica. Trattandosi di strumenti e di attori molto diversi tra loro, occorre individuare la tecnica di comunicazione maggiormente efficace onde evitare che vengano travisati i risultati ottenuti o, ancor peggio, che vengano prodotte e generate fake news. La principale distinzione è fra stakeholder e pubblico generalista e per ciascuna di queste categorie verranno presentate le strategie di public engagement e le tecniche redazionali più adatte per un’efficace comunicazione della ricerca.