



Unione europea
Fondo sociale europeo



INVITO A PRESENTARE PROGETTI DI FORMAZIONE ALLA RICERCA IN ATTUAZIONE DEL
PIANO TRIENNALE ALTE COMPETENZE PER LA RICERCA, IL TRASFERIMENTO
TECNOLOGICO E L'IMPRENDITORIALITÀ APPROVATO CON DELIBERAZIONE
DELL'ASSEMBLEA LEGISLATIVA N. 38 DEL 20/10/2015
POR FSE 2014/2020 Obiettivo tematico 10"

Delibera di Giunta regionale n. 388 del 19/03/2018
Scadenza 26/04/2018

scheda descrittiva dei Progetti di formazione alla ricerca

Titolo del progetto di formazione alla ricerca

Multimedia Data Learning: Big Multimedia Data Processing, Understanding, Retrieval
by Learning

(selezionare uno dei seguenti ambiti)

- Ambito A) "RISORSE UMANE PER UN'ECONOMIA DIGITALE: BIG DATA"
 Ambito B) "RISORSE UMANE PER LA SPECIALIZZAZIONE INTELLIGENTE"
 Ambito C) "RISORSE UMANE PER IL PATRIMONIO CULTURALE"

Ateneo

Università di Modena e Reggio Emilia

Corso di dottorato e relativo il ciclo

Dottorato in ICT- XXXIV ciclo

Coordinatore del Corso di dottorato

Sonia Bergamaschi

Referente scientifico/supervisore
borsa di dottorato:

Rita Cucchiara

Telefono:
335 7182399

e-mail:
Rita.cucchiara@unimore.it

Dipartimento

Dipartimento di Ingegneria Enzo Ferrari

Sintesi del progetto di formazione alla ricerca

L'analisi di grandi archivi multimediali (immagini, video, testi digitalizzati, audio, dati 3D) e' uno dei temi portanti dell'ICT nell'ambito del "**multimedia e multimodal processing**", che da più di un decennio ha portato enormi avanzamenti tecnologici e risultati tangibili: i sistemi di ricerca di Google, Youtube, Facebook ne sono l'esempio più concreto: Le stesse tecnologie sono ora impiegate dai broadcaster multimediali (RAI, SKY..) ma ultimamente si stanno diffondendo anche in modo più capillare, considerando la facile raccolta di big data multimediali anche da aziende di piccole- medie dimensioni e di tutti i settori merceologici.

Le tecnologie in questo campo hanno subito una tremenda rivoluzione con la introduzione di modelli di **Deep Learning**: se prima i dati MM erano rappresentati da "feature vector" selezionati manualmente (istogrammi colore, vettori MPEG etc per dati visuali, ontologie testuali, frequenze per audio) ora lo sono attraverso descrittori neurali sia da architetture generiche opensource sia ad hoc per i vari archivi. La comprensione del contenuto, la ricerca ed il recupero anche **cross-modale** (ad es. da "captions" testuali ad immagini o viceversa), il confronto tra video per definire copie o copie parziali, la comprensione del contenuto di componenti digitalizzati hanno ora metodologie completamente nuove basate **sull'intelligenza artificiale sub-simbolica, sul machine learning e su metodologie nuove di computer vision**. Questi temi di ricerca di base sono lo stato dell'arte nel multimedia e possono essere affrontati solo con strategie a medio termine, grazie a percorsi di dottorati di ricerca e dalla concomitanza di tre fattori

- a) disponibilità di big data multimediali (sia "stored" che "streamed")
- b) competenze di frontiera sul deep learning
- c) disponibilità ed uso consapevole di supercalcolo grafico ossia di server GPU.

Questo e' possibile ad UNIMORE AimageLab anche grazie alla disponibilità di server di calcolo dedicati e alla stretta collaborazione con CINECA. Il progetto, portato avanti da almeno un dottorando di ricerca, in collaborazione con altri ricercatori, vuole esplorare nuovi algoritmi e tecnologia di frontiera nei sottotemi collegati di a) Video content understanding, b) Cross-modality retrieval (da testo a immagini e video e 3D e viceversa) c) Copy and similarity detection.

Il progetto di dottorato prevede l'impegno di dataset generici open-source impiegati in tutti i centri di ricerca multimediale (come COCO Microsoft, FLikr8000) altri specifici di video broadcast/social (provenienti da Facebook, RAI e SKY) altri specifici di dominio (es. immagini di abbigliamento, video da auto in movimento da aziende automotive, video urbani da telecamere di sorveglianza, immagini artistiche/storiche proveniente dalla digitalizzazione di archivi). Prevede lo studio di nuovi modelli di Deep Learning generalizzabili cross- dominio (tra cui l'uso di reti CNN siamesi LSTM, GRU, GANS, con modelli di metric-learning, active-learning e reinforcement-learning), ridefiniti e riaddestrati dall'inizio sui dati anche mediante tecniche di "augmentation" e di "privileged information" date dalla supervisione umana, lo studio di convergenza, di robustezza e di generalizzazione, la creazione di tools di sviluppo e di applicazioni concrete. Infine la definizione di nuovi approcci da proporre nelle più importanti sedi internazionali, anche in collaborazione con centri di ricerca esteri per portare la ricerca sui big data del nostro territorio a livelli di eccellenza internazionale.

Finalità generali e i risultati attesi del progetto di formazione alla ricerca

Finalità:

- Studio di il ciclo di vita di estrazione di conoscenza da sistemi multimediali delle prossime generazioni con nuovi modelli di elaborazione (processing), comprensione (understanding), riconoscimento semantico di similarità (matching, detection), ricerca e recupero (search and retrieval) di dati multimediali
- Realizzazione di sistemi allo stato dell'arte sul DL per la Autonomous and Augmented Intelligence, sia di tipo completamente automatico sia di tipo interattivo con l'essere umano (come modelli di active-learning)
- Studio approfondito di modelli innovativi di DL per il multimedia, di tipo generale, delle prestazioni, della robustezza e capacità di generalizzazione
- Definizione di metodologie di progettazione di tools di apprendimento automatico sui big MM data specializzabili in diversi ambiti applicativi

Risultati attesi:

- nuovi modelli di Intelligenza Artificiale per i dati multimediali;
- nuovi tools di apprendimento per la comprensione, il retrieval ed il copy detection;
- specializzazioni verticali in almeno due campi diversi, con le stesse metodologie quali il campo dell'abbigliamento, dell'automotive, del broadcasting e delle smart cities.
- Realizzazione di articoli scientifici internazionali sulle conferenze e journal di fascia A (ACM MM, CVPR, ICCV, NIPS; T-MM, T_PAMI)
- Realizzazione di brevetti
- Realizzazione di progetti Europei ed internazionali con centri di ricerca ed aziende del territorio
- Eccellenza del territorio Emiliano Romagnolo nei big data ed aumento della attrattività produttiva

Coerenza del progetto rispetto alla Strategia di specializzazione intelligente e alle linee programmatiche di sviluppo regionale

Il progetto e' perfettamente coerente alle strategie regionali nell'ambito -Dei Big Data nell'ambito "BIG DATA IN ICT AND DIGITAL CONTENT" come strategie generali applicabili anche in altri contesti applicativi quali "SOCIAL SCIENCE AND HUMANITIES" e "SMART, CITIES, SAFETY & SECURITY" e con diretti collegamenti anche al mondo delle "HUMAN BRAIN AND NEUROSCIENCE COMPUTING"

Il progetto e' inoltre perfettamente coerente con le S3 nell'ambito -Innovazione dei servizi ed Industrie culturali e creative

E' infine coerente con tutte le politiche Regionali di specializzazione intelligente verso il rafforzamento dell'ecosistema regionale, della eccellenza territoriale, della apertura all'innovazione con contatti internazionali.

Per i progetti di cui all'ambito B) descrivere la focalizzazione rispetto alle value chain più rilevanti per l'economia regionale anche coerentemente con gli ambiti di attività dei Clust-ER

Il progetto nasce specificatamente per lo sviluppo di tecnologie di big data A) ma e' inoltre perfettamente coerente con le S3 nell'ambito
-Innovazione dei servizi: value chain "Intelligent IT Services", per lo sviluppo di servizi di recupero e comprensione di dati multimediali
- Industrie culturali e creative : value chain "multimodal e nuovi modelli di business" anche epr le enormi potenziali applicazioni anche con PMI e nuove startup nell'ambito multimediale

Per le potenziali applicazioni e' anche in stretto contatto con altre strategie che impiegano big data processing, nell'automotive, nella motoristica, nel tessile e nel medical imaging.

Conoscenze e competenze attese e descrizione della loro declinazione nel sistema economico produttivo

Si prevede che un dottorato di ricerca ottenga dal percorso di studi le competenze e conoscenze attualmente più spendibili nell'ambito dell'ICT ossia conoscenze nello sviluppo di sistemi e servizi di AI di Deep Learning e Computer Vision, di sistemi di video analytics e di retrieval.

Il dottorato nasce con collegamenti internazionali e fornirà competenze di frontiera in campo scientifico-tecnologico dell'ICT declinate in tutti i settori emergenti dell'IT e dei big data ma anche a supporto del sistema produttivo manifatturiero, tessile e automotive.

Queste competenze diventano sempre più richieste anche in ambito medico e di e-health essendo assolutamente trasversali e pervasive.

L'obiettivo e' formare nuove figure estremamente specializzate in campo Artificial Intelligence ed Information Technology e nello specifico campo del Machine (Deep) Learning su Big MM Data.

Spendibilità nel sistema economico produttivo delle conoscenze e competenze e analisi dei risultati occupazionali attesi

Queste competenze sono tra le più richieste nel mondo secondo i dati del 2018 di Tractica e di Gartner con un incremento del mercato superiore al 20% nei prossimi 5 anni in tutto Europa. Tractica addirittura stima la revenue dell'AI nel mondo in crescita da \$3.2 billion nel 2016 a \$89.8 billion by 2025.

<https://www.tractica.com/research/artificial-intelligence-market-forecasts/>

Sono ora competenze introvabili ed assai richieste nei grandi player internazionali (IBM, Microsoft, Amazon, Facebook, Google), dalle aziende manifatturiere locali che lavorando su archivi multimediali (aziende automotive e tessili e di robotica collaborativa) e soprattutto da aziende IT.

Sono assolutamente spendibili anche nella creazione di nuove imprese come startup innovativa e spinoff sull'AI.

Da un recente articolo dle NewYork Time si evince come questo problema sia sentito in tutto il mondo High Tech (<https://www.nytimes.com/2017/10/22/technology/artificial-intelligence-experts-salaries.html>) e certamente anche in Italia e nella nostra regione che si sta caratterizzando per uno dei piu' alti numeri di startup innovative nell'AI.

Contestualizzazione del progetto: descrizione delle iniziative di ricerca e innovazione nelle quali si colloca la proposta

Il progetto si pone nell'ambito di una ricerca sul multimedia, sull'AI e sulla visione che da più di 10 anni il laboratorio di ricerca AImageLab di UNIMORE sta svolgendo.

La ricerca nasce nell'ambito di iniziative internazionali quali

- Collaborazione con Laboratorio Facebook AI Reserach- Paris sulla video copy detection, anche attraverso il grant ricevuto nel 2017 da AImageLab che si colloca tra i 15 lab FAIR Europei; in questo ambito si stanno studiando nuovi emtodi di video similarity matching, impiegabili in ogni contesto applicativo;
- Collaborazione con Hebrew University (Prof. Naphtaly Thisby) sulla teoria della convergenza delle reti neurali in ambito di information theory; grazie a questa collaborazione prluriennale si vologliono esplorare anche aspetti teorici dei meccanismi neuronali;
- Collaborazione con Technion Univesrity (prof. Lihi Zelik-Manor) nell'ambito di ricerca attentzionale e di salienza nelle immagini anche per analisi di video comportamentali.

Grazie a queste iniziative di ricerca gia' in atto al presenza stabile di uno studente di dottorato permetetra' di rendere piu' stabile la collaborazione, di formalizzare accordi e scambi di ricerca a lungo termine.

Esso si colloca anche nelle iniziative di innovazione in collaborazione con aziende del terriorio quali Yoox per il retrieval in archivi di abbigliamento, Maserati per la comprensione di video di autonomous driving, Ferrari per la comprensione della attenzione del guidatore (lab RedVision), IBM per la comprensione di big data multimediale ed aziende nazionali quali Sky con cui sono gia' attivi contatti.

Infine il progetto si inserisce nelle iniziative di ricerca legate ai Cluster regionali della "innovazione dei servizi" per le tecnologie IT sull'intelligenza artificiale e della nascente associaizione Regionale sui Big Data.

Contestualizzazione del progetto: descrizione dei progetti competitivi maggiormente rilevanti nei quali si colloca la proposta

- In questo momento sono attivi alcuni progetti competitivi con cui il potenziale dottorando si dovrà confrontare quali
- CultMedia (progetto MIUR CLuster Tecnologie beni culturali) " comprensione e riuso di video culturali"
- Prystine (progetto EU Horizon2020 sull'autonomous driving, in cui Aimagelab ha a carico il task di comprensione di video da auto in movimento)
- Sono inoltre attive alcune proposte europee in fase di valutazione quali LeHI (passato il primo step di valutazione su CHIST-ERA) per riconoscimnto di big dati visuali per robot sociali in citta' e Souvenir (in valutazione) sul cultural heritage

Contestualizzazione del progetto: descrizione delle collaborazioni con soggetti pubblici e/o privati a livello nazionale e internazionale nelle quali si colloca la proposta

Aimagelab ha attualmente attive collaborazioni pertinenti quali

- Collaborazione con Panasonic Silicon Valley Lab per supporto a dottorati e scambi di studenti con E.Francisco e Giappone disponibilità di big data
- Collaborazione con Facebook AI Research Lab
- Collaborazione con RAI nel progetto "città Educante" Miur
- Collaborazione con Comune di Modena e Maserati nel progetto MASA per l'analisi di big data da telecamere fisse ed in movimento
- Collaborazione ISCRA con CINECA per uso di supercomputer per big data;
- Collaborazione con NVIDIA AI City (Dr. Milhind Naphade) California per riconsocimeento di big data multimediale di auto in movimento
- Collaborazione con DIVA IARPA Project Usa per la raccolta di big data sullel azione dei cittadini e la loro interpretazione automatica. UNIMORE e' attualmente sottocontraente dell'azienda Kittware che partecipa a DIVA
- Collaborazione con University di Amsterdam (Prof. Cees Snoek) che collabora in DIVa e che gestisc ei laboratori Qualcomm sul multimedia.

Descrizione delle modalità con cui il beneficiario della borsa potrà essere coinvolto nelle iniziative/progetti/collaborazioni

- Collaborazione diretta con i centri di ricerca con impiego di server GPU e di risorse di calcolo in comune;
- Collaborazione con stakeholder industriali (e broadcaster) per impiego di big data multimediali a disposizione e supporto alla raccolta ed alla annotazione semi-automatica
- Collaborazione con scambi internazionali (si prevede obbligatoriamente almeno 6 mesi all'estero)
- Collaborazione nei progetti EUB esistenti e nella presentazione di nuovi progetti

Descrizione delle eventuali collaborazioni con soggetti pubblici e/o privati che si intendono attivare per la realizzazione del progetto, da formalizzare prima dell'avvio del ciclo di dottorato di riferimento, indicando ruolo e contributo apportato, e accordi relativi alla proprietà intellettuale

Nessuno.

Le collaborazioni indicate (Facebook, Panasonic, Ferrari, Maserati e Comune di Modena) sono già in essere ed altre (IBM) sono in fase di formalizzazione; già prevedono accordi di proprietà intellettuale e non e' necessario formalizzare altri accordi. Le attività all'interno dell'Associazione Big Data Regionale sarà normata dagli accordi opportuni.

Descrizione delle ricadute attese sul sistema regionale dell'innovazione e della ricerca nella prospettiva e nel contesto nazionale e internazionale

Le ricadute attese sono moltissime:

- Aumento della visibilità scientifica internazionale della Regione Emilia Romagna in uno dei campi di ricerca più promettenti dei prossimi anni
- Aumento delle relazioni internazionali stabili, possibilmente nell'ambito della Associazione Big Data Regionale, dei cLust-ER e delle università regionali
- Miglioramento delle capacità tecnologiche delle aziende del territorio in particolare delle aziende IT e delle startup, ricevendo dal progetto informazioni, dati e modelli open-source e formazione continua
- Miglioramento del potenziale produttivo delle aziende manifatturiere del territorio per imparare l'uso di Deep Learning, AI applicata ai dati multimediali e risolvere problemi reali nella gestione degli archivi multimediali a supporto dei processi produttivi
- Aumento del potenziale di innovazione regionale nella gestione dei dati multimediali, per la creazione di nuove imprenditorie, nuovi posti di lavoro e per il rafforzamento dell'economia nazionale in ambito AI.

Si ricorda a questo proposito che il recente "libro bianco dell'Intelligenza Artificiale" presentato dall'AGID in marzo 2018 colloca l'emilia Romagna ai primi posti in termini di aziende che lavorano in ambito AI.

La possibilità di avere uno o più studenti di dottorato in quest'area e quindi esperti nel settore con competenze allo stato dell'arte internazionale ne favorirà tutto il contesto produttivo nella prossima decade.