

## MMSP LABORATORY WORKSHOP

**Tuesday February 16, 2021**  
**on Webex**

**Sofia Bellotti, Gabriele De Mattia, Cesare Bagnolini, Fabio Cimmino, Andrea Gusmara and Alessandro Di Crescenzo,**

Students at the MMSP laboratory  
will present their thesis works

**08:30 – Francesca Gasparini**

**Introduction**

**08:45 – Sofia Bellotti**

**The MAMI project: designing web architecture to disclose the problem of online misogyny**

L'intero lavoro di tesi è stato svolto a sostegno del progetto MAMI, nato all'interno dell'Università degli Studi di Milano Bicocca, finalizzato alla progettazione di algoritmi in grado di identificare automaticamente contenuti misogini.

Partendo dall'analisi del fenomeno della misoginia in contesti online, quali sono le sue radici, come si manifesta, perché e con quali conseguenze, si delinea la criticità e la necessità di intervento.

Vengono quindi presentati alcuni algoritmi progettati per l'identificazione di contenuti aggressivi e/o misogini online già presenti nella letteratura scientifica. Da questa analisi dello stato dell'arte si decreta l'unicità e l'originalità del lavoro svolto dal team del MAMI project, soprattutto per quanto riguarda il focus sui meme, una tipologia di contenuto sempre più dilagante nel web.

Si denota pertanto la forte necessità di ampliare un progetto così innovativo, aprendolo al pubblico di esperti e non esperti. Nasce così il sito web The MAMI project.

Sviluppata con una metodologia User Centered, la piattaforma si propone di essere uno strumento di divulgazione e ricerca, abile anche nel coinvolgere l'utente a diventare parte attiva del progetto.

Dopo un'analisi della progettazione in tutte le sue fasi, a partire dalla ricerca, passando per la valutazione tramite user test, fino ad arrivare allo sviluppo e implementazione, la piattaforma web The MAMI project prende vita.

L'elaborato si conclude con uno sguardo sul futuro, perché questo progetto è in grado di acquisire una rilevanza sempre maggiore quanto più verrà utilizzato e condiviso.

#### **09:15 – Gabriele De Mattia**

##### **Piattaforma MAMI: sviluppo di una web application per la raccolta e classificazione di meme misogini**

Il mio lavoro di tesi è finalizzato alla collaborazione al progetto di ricerca MAMI, sviluppato da un gruppo di ricercatori dell'Università degli Studi di Milano-Bicocca. Nello specifico, il lavoro da me realizzato ed illustrato nel mio elaborato è finalizzato ad una raccolta di valutazioni a immagini fornite da utenti volontari, le quali saranno utili per alimentare la banca dati di una intelligenza artificiale, che potrà procedere con la valutazione autonoma sulla presenza di elementi misogini all'interno di uno degli strumenti di comunicazione più diffusi sul web: il meme. Ho progettato e realizzato il database per la raccolta dei dati che alimenteranno l'intelligenza artificiale, attraverso la web application che ho sviluppato per permettere agli utenti di fornire il loro contributo e ai ricercatori di gestire i dati raccolti, il tutto integrato alla perfezione con il sito web del progetto, realizzato da una studentessa di un altro corso di laurea.

#### **09:45 – Cesare Bagnolini**

##### **Applicazioni di tecniche di OCR su contenuti memetici**

Si utilizza Tesseract, un software OCR in grado di riconoscere il testo presente in un'immagine e restituirlo come stringa, per estrarre il contenuto di meme misogini.

Si definisce una metrica full reference per valutare la qualità del suo output quando si è in possesso della ground truth, ovvero il reale contenuto dei meme.

Viene definita una pipeline generica per estrarre il testo dei meme che comprende una fase di pre-processing dell'immagine, l'estrazione realizzata con Tesseract e una fase di pulizia del testo estratto.

Variando la tecnica di elaborazione dell'immagine utilizzata si definiscono nuove pipeline.

Dato che in situazioni reali non si è in possesso della ground truth, si definisce una metrica no reference per valutare la qualità dell'output dell'OCR.

Applicando questa metrica no reference si individua quale pipeline sia la più conveniente da usare.

Si dimostra la correlazione tra i risultati della metrica full reference e quelli della metrica no reference, rispetto agli output della pipeline più conveniente, attraverso un test statistico che indica la correttezza della metrica no reference.

#### **10:15 – Coffee break**

**10:30 – Fabio Cimmino**

**Deep learning strategies for motor imagery electroencephalographic signals – from raw to time frequency representation**

Electroencephalography (EEG) allows to measure the electrical activity produced by the brain and it proved to be an important tool for neuroscientific applications, such as Brain-Computer Interface.

However, decoding EEG signals is still a current challenge given the intrinsic complexity of the signal, limiting the usefulness of EEG-based Brain-Computer Interface applications. For this reason, Deep Learning techniques are used to improve EEG signal processing and are showing great promise with their high generalization and expressive feature extraction capabilities compared to traditional methods.

By developing and testing different inputs (from raw to time-frequency representation) and architecture formulations of two Deep Learning networks present in the literature, this thesis aims to understand how much similarity there is between EEG-based real movement and EEG-based imaginary movement.

The work focused on the movement of the left and right hand. It has been observed that, for a Deep Learning classifier, the left/right hand real movement is hardly discernible from the left/right hand imaginary movement.

Therefore left/right-hand real movement can be replaced by the corresponding imaginary movement in order to exploit motor imagery tasks in Brain-Computer Interface systems.

**11:00 – Andrea Gusmara**

**Analisi dati inerziali e fisiologici acquisiti con sensori indossabili**

Ormai da tempo in tutte le popolazioni del mondo si sta presentando un costante incremento del numero di anziani, per lo più persone che generalmente perdono dimestichezza in attività che prima sembravano semplici, come per esempio una camminata.

A questo proposito propongo un esperimento in cui sono coinvolti la Gait Analysis e l'Inertial Pedestrian Navigation System. Aspetti inseribili all'interno dei progetti di Walkability e di Health Care Monitoring.

Per quanto riguarda la Gait Analysis mi concentrerò sulla valutazione e il calcolo dei parametri ricavabili dai soli dati inerziali.

Nella prima fase di analisi andrò a concentrarmi sull'individuazione del passo (StepDetection).

Dopodiché andrò a valutare i seguenti parametri: Stride Length, Step frequency, Stride velocity. Alcuni di questi ultimi parametri saranno successivamente molto utili anche nella fase di calcolo della navigazione inerziale del soggetto durante la camminata.

Per quanto riguarda il Inertial Pedestrian Dead Reckoning, andrò a utilizzare l'orientamento e la lunghezza del passo, in modo tale da ricavare lo spostamento.

L'esperimento su cui si baserà la mia analisi consiste sostanzialmente in una camminata lungo un percorso prestabilito e vincolato ad una frequenza di passo.

Il sensore scelto per effettuare le acquisizioni dei dati inerziali è il sensore Shimmer 9DOF.

L'obiettivo di questa tesi è quello di acquisire dei dati inerziali e di velocità angolare validi e interessanti per effettuare una buona analisi sperimentale di Position Dead Reckoning e di Gait analysis, riuscire a ricostruire il tracciamento del percorso con buona qualità e precisione basandomi sull'orientamento e la lunghezza del passo, valutare i risultati sulla stima dei parametri legati alla Gait Analysis.

**11:30 – Alessandro Di Crescenzo**

**Picture-in-picture e Voice-over: uno studio sperimentale su due tipologie di DaD a confronto durante l'emergenza da COVID-19.**

A causa della pandemia da Covid-19, l'ultimo anno accademico è stato caratterizzato da una totale erogazione di contenuti didattici a distanza. Questo lavoro di tesi si inserisce in tale contesto, analizzando sperimentalmente due tipologie di video-lezioni: Picture-in-picture (PIP) e Voice-over (VO).

In primo luogo, viene presentato lo stato dell'arte, relativo alla risposta di diversi paesi del mondo costretti alla DaD a partire da marzo 2020, con particolare attenzione all'Italia e alla Cina. Successivamente il cerchio si stringe sul nostro Ateneo: vengono infatti somministrati dei questionari a docenti e studenti Bicocca, i cui esiti hanno guidato l'elaborazione del setting sperimentale. Quest'ultimo è consistito nella somministrazione di una video-lezione costituita da quattro moduli didattici sulla Fotografia Digitale, creati sia in versione PIP che VO. I soggetti sperimentali sono stati studenti universitari che, nell'ultimo anno, avevano seguito lezioni a distanza. Ogni partecipante ha visto un video con metà lezione in versione PIP e un video con l'altra metà in versione VO. Dopo la visione di ciascun video, i partecipanti hanno risposto a un quiz di venti domande. È stato ipotizzato che la video-lezione in modalità PIP, preferita dagli studenti, avesse maggiore efficacia, quantificata dai punteggi ottenuti, e che il primo modulo di ogni coppia presentata fosse quello in cui gli studenti eccellessero maggiormente, essendo meno stanchi. I risultati sembrano confutare entrambe le ipotesi, in quanto non sono emerse significatività statistiche, sebbene le medie dei punteggi dei video in modalità PIP siano sempre risultate maggiori della versione VO.

L'elaborato si conclude con uno sguardo al futuro: valutare l'efficacia della DaD tramite eye-tracking via webcam, di cui viene presentato un prototipo.