Who Gets Heard? Investigating Accent and Gender Disparities in Automatic Speech Recognition

Tisha Prasad, STEM4Change

Automatic speech recognition (ASR) systems have achieved remarkable performance in recent years, yet concerns remain regarding their robustness across speakers with diverse linguistic backgrounds and gender identities. Differences in pronunciation caused by non-native accents, alongside variation in speech characteristics across gender, can lead to unequal recognition performance and introduce potential fairness and quality-of-service issues. This study evaluates the Whisper-tiny.en ASR model using the GMU Speech Accent Archive, measuring ASR performance across native and non-native English speakers from over 150 accent groups, with a focus on how gendered performance differences manifest across accents. Performance differences are quantified using Word Error Rate, Character Error Rate, Match Error Rate, and Word Information Lost. The results reveal substantial disparities across accents, with East and Southeast Asian language backgrounds (e.g., Mandarin, Cantonese, Japanese, Korean, Vietnamese, Lao), as well as a subset of African language backgrounds (e.g., Somali, Bai, Miskito, Fang) exhibiting markedly higher error rates than speakers with English and European language backgrounds, (e.g. Afrikaans, Dutch, German, Swedish, Hungarian). Across all accents, female speakers achieve slightly lower error rates than male speakers, suggesting that while accent is the primary source of performance variation, gender-related differences persist.

COMPARATIVE ANALYSIS OF FEATURE EXTRACTION APPROACHES IN A SINGLE-CHANNEL MULTISENSORY SETTING

Fouziah Md Yassin, Universiti Malaysia Sabah

Motor imagery-based brain-computer interfaces (MI-BCIs) commonly rely on multichannel electroencephalography (EEG) to extract discriminative neural information. However, the complexity of multichannel systems limits their practicality for everyday applications. Single-channel EEG offers a more practical alternative but presents challenges in extracting informative features due to reduced spatial information. This study compares different feature extraction approaches for single-channel prefrontal EEG recorded from AF3 and AF4 during motor imagery tasks performed under multisensory conditions. The evaluated approaches include time-domain features, time-frequency features (discrete wavelet transform (DWT)), a fusion of both feature sets, and a features fusion with feature selection. Feature selection was performed using correlation analysis and the Kruskal–Wallis's test to retain the most discriminative features before classification. The approaches were evaluated using EEG data collected from healthy participants performing left-right and forward-backward MI tasks under standard, visual, auditory, and olfactory stimuli. The comparative analysis demonstrates that feature fusion with feature selection generally improved binary classification performance across classifiers compared with individual feature extraction approaches.

Bridging Data Silos: A Two-Step Unsupervised Approach to Record Linkage

Zaturrawiah Ali Omar, Universiti Malaysia Sabah

Linking records across databases, whether to identify duplicates or match the same entity across sources, is a foundational step in producing trustworthy, connected data. The most established approach, probabilistic record linkage, still depends on carefully set similarity thresholds, a process that becomes difficult without access to labelled "gold standard" data to validate against. Supervised machine learning approaches can offer an alternative, but they introduce their own dependency on labelled training data, which is often expensive or simply unavailable. This talk presents a two step approach built on the Unsupervised Random Forest (URF) algorithm, which generates a similarity score for record pairs without requiring any labels at all. We explore how this single algorithm can serve two purposes: detecting duplicate records directly, and automatically selecting high quality training data to support supervised classification, offering a path toward record linkage that is less reliant on labelled data.

Mining the Foundations of Truth: Extracting Structured Knowledge Graphs from Unstructured Classical Texts

Norhafiza Hamzah, UMS

As the volume of digital information grows, global artificial intelligence systems increasingly struggle to process highly specialized, culturally nuanced data without hallucinating. Aligning with the IDWSDS 2026 theme of "Creating Global Connections", this talk frames domain-specific fact-checking fundamentally as a complex data mining problem. Using a corpus of Islamic jurisprudence (Hadiths) as a primary case study, we explore the methodological challenges of mining unstructured, classical texts to discover and extract structured data representations. The presentation will detail the end-to-end data mining pipeline, including NLP-driven entity recognition, coreference resolution, and relationship extraction, required to transform raw text into a mathematically traversable Knowledge Graph. By demonstrating how to mine and structure these deeply embedded contextual rules, this talk illustrates how rigorous data mining serves as the critical first step in enabling accurate, culturally resilient automated verification systems on a global scale.