**Introduction**

- **Motivation**
  - Computing devices have been easily accessible and information search has been a common part of regular conversations, where these meetings include discussions for identifying participants’ next actions.

- **Main Idea**
  - Human-machine interactions collected by existing intelligent systems (e.g., Cortana data) may help detect actionable items in human-machine dialogues (e.g. meetings)
  - Learning action representations using a CDSSM architecture helps transfer high-level semantics across genres

- **Actionable Item Detection Task**
  - Goal: provide the easy access to information and perform actions a personal assistant can handle without interrupting the meetings
  - Assumption: some actions and associated arguments can be shared across genres

---

**Convolutional Deep Structured Semantic Models (CDSSM)**

- **Model Architecture**
  - **Semantic Layer:** γ
    - Projection Matrix: W
  - **Max Pooling Layer:** \( i_\gamma \)
    - only retain the most prominent local features by applying the max operation over each dimension of, to keep the max activation of hidden topics across the whole word sequence
  - **Convolutional Layer:** \( i \)
    - contextual features \( c_i \), for each target word \( \tilde{w}_i = tanh(W(c_i)) \)
  - **Word Hashing Layer:** \( L \)
    - one-hot word vector \( \rightarrow \) tri-letter vector
      - e.g. “email” \( \rightarrow \) “eml”, “ema”, “mai”, “a”, “m”
  - **Word Sequence:** \( x \)
    - user utterance / intent

- **Training Procedure**
  - **Predictive model:**
    - \( \Lambda(\theta_1) = \log \prod P(U^+ | U) \)
  - **Generative model:**
    - \( \Lambda(\theta_2) = \log \prod P(U^+ | A) \)

- **Bidirectional Score Estimation**
  - incorporate the effectiveness of predictive and generative models
  - \( S_{Bi}(U, I) = \gamma S_P(U, I) + (1-\gamma)S_G(U, I) \)

- **Conclusion**
  - The latent semantic features generated by CDSSM show the effectiveness of detecting actions in meetings compared to lexical features, and also outperform the state-of-the-art semantic features.
  - The adaptation techniques are proposed to adjust the learned embeddings to fit the target genre when the source genre does not match well with target genre, showing significant improvements in detecting actionable items.
  - The paper highlights a future research direction by releasing an annotated dataset and the trained embeddings for actionable item detection.