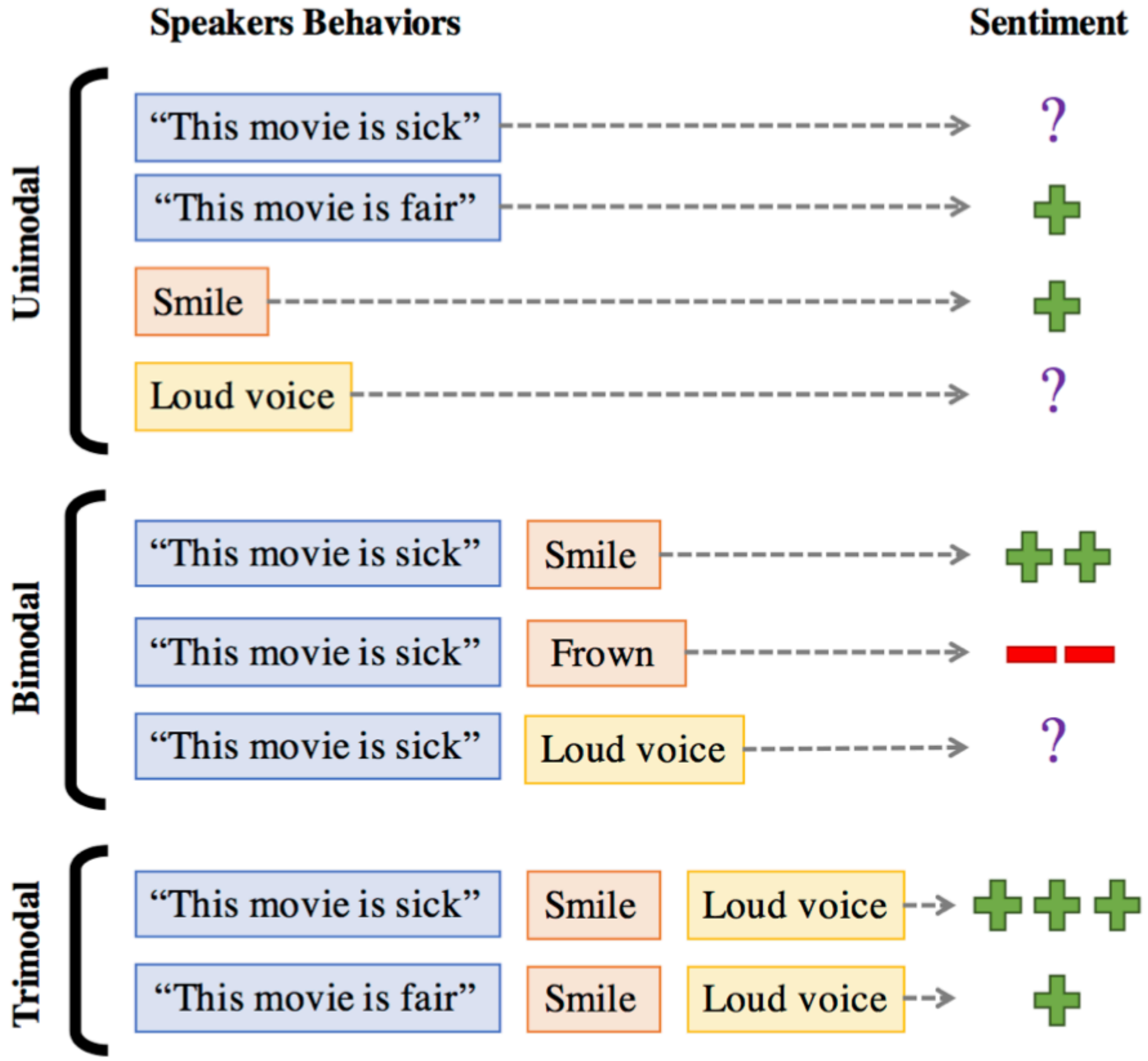
Privacy-Preserving Multimodal Sentiment Analysis

Introduction

With the proliferation of social media, the importance of multimodal sentiment analysis has attracted the attention of researchers for stock market performance prediction, election outcome prediction, customer satisfaction assessment and brand perception analysis.

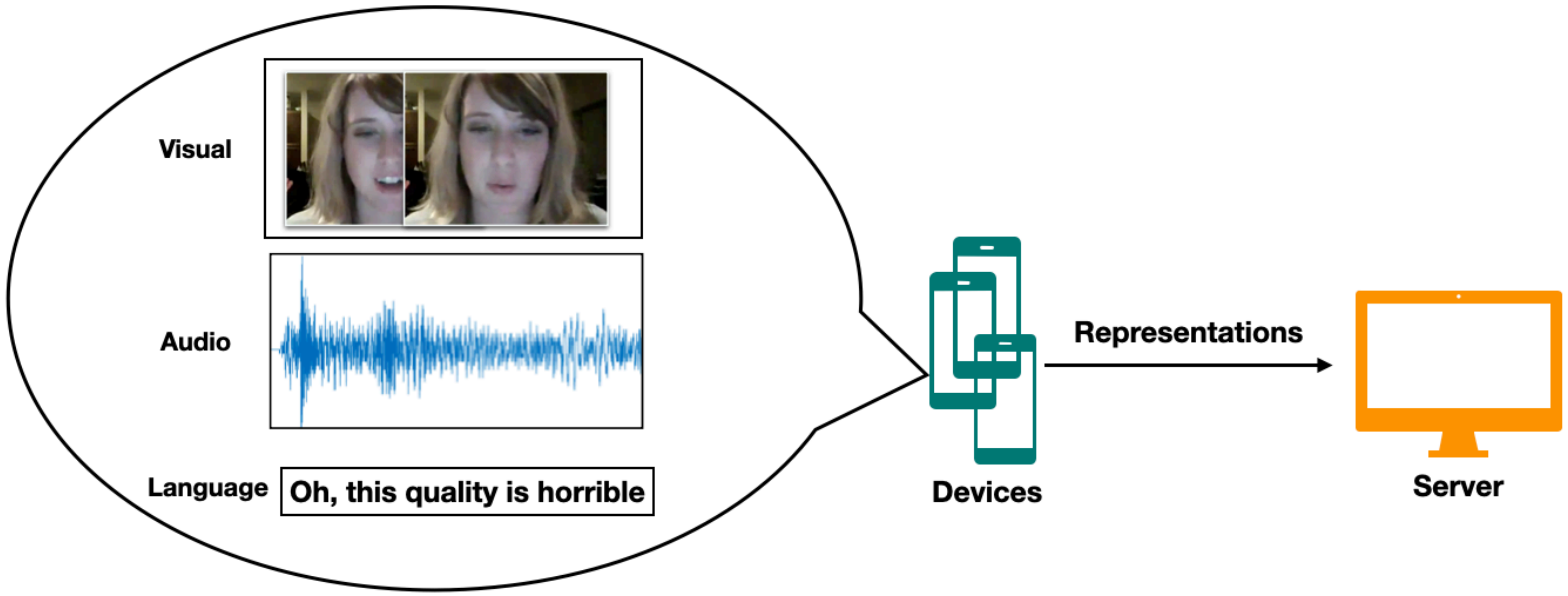


Privacy-Preserving Multimodal Sentiment Analysis

Introduction

Driven by the explosive progress of deep learning technology, learning-based prediction has been treated as one promising and effective approach to realize multimodal sentiment analysis through multimodal data representations extracted from raw multimedia data.

Unfortunately, the extracted data representations can be exploited to infer private information by malicious attackers, causing serious privacy threats and substantial economic loss to individuals.



Related Work

i) Adversarial Training-Based Models

ii) Differential Privacy-Based Approaches

iii) Differentially Private Transform-Based Methods

Problems:

i) The adversarial training-based models cannot ensure a privacy protection guarantee.

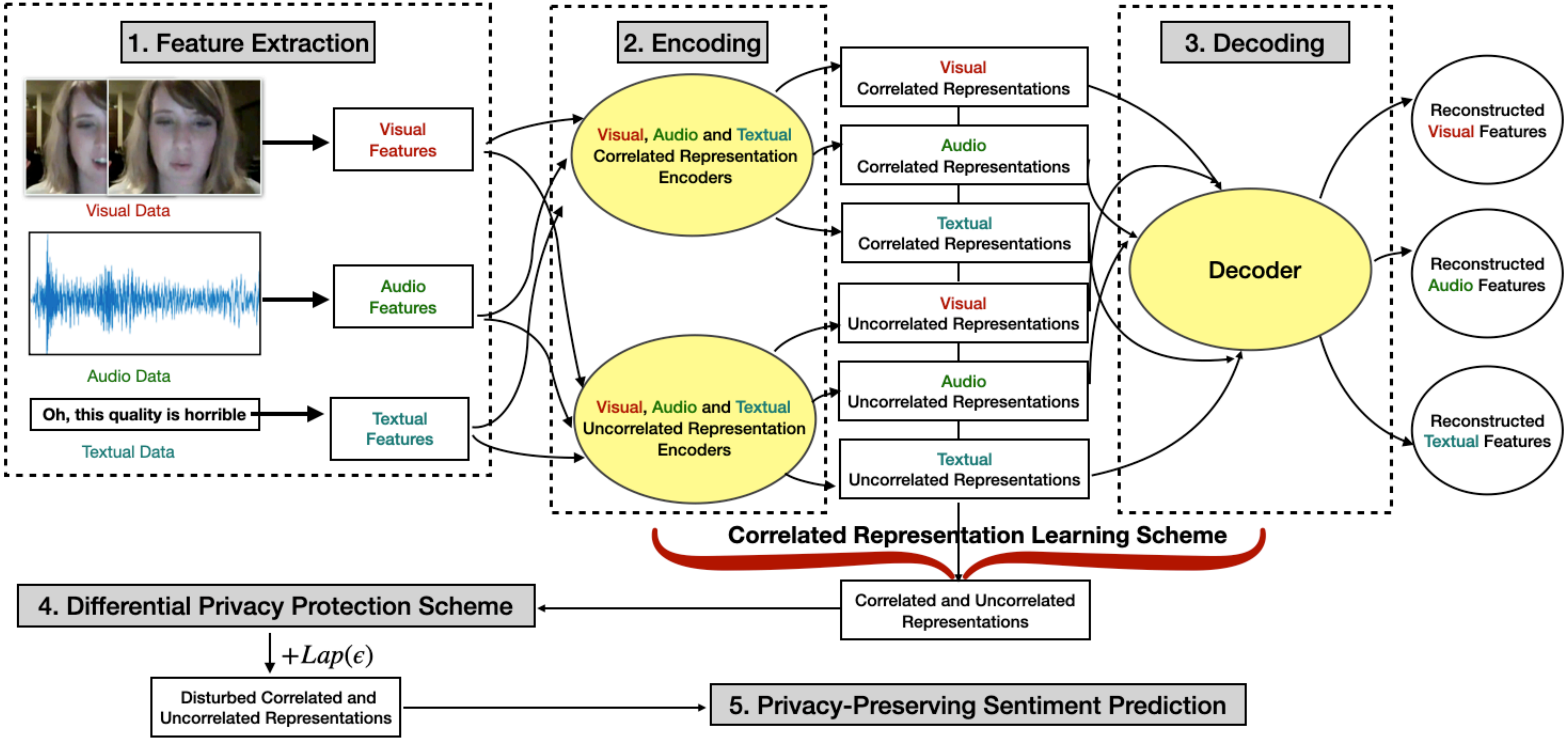
ii) For correlated data, the added Laplace noise should be increased with the growth of data correlation, which sacrifices the performance of learning models.

iii) The existing transform-based methods can only be exploited to transform the low-dimension data into an independent data domain and thus cannot be applied to the high-dimension multimodal data.

Challenge:

It is a challenging task to generate the privacy-preserving representations of high-dimension correlated multimodal data without reducing the performance of multimodal sentiment analysis.

Differentially Private Correlated Representation Learning (DPCRL)



1. Feature Extraction

The stacked bi-directional Long Short-Term Memory scheme (sLTSM) is exploited to map multimodal data into a feature vector:

$$\mathbf{f}_m = sLSTM(\mathbf{U}_m; \theta_m^{slstm})$$

2. Encoding

For each feature vector, its correlated and uncorrelated representations should capture two distinctive aspects of the same modality data.

Any two of the uncorrelated representations should be distinctive without redundancy.

The correlation between any two of the correlated representations should be close to the correlation factor as much as possible.

So, we use the correlated multimodal representation encoder to extract the correlated representation and use the uncorrelated multimodal representation encoder to capture the uncorrelated representations:

$$\mathbf{f}_m^c = E_m^c(\mathbf{f}_m; \theta_m^c, c),$$

$$\mathbf{f}_m^u = E_m^u(\mathbf{f}_m; \theta_m^u),$$

2. Encoding

We formulate the data orthogonality loss:

$$\mathcal{L}_{enc_1} = \sum_{m \in \{v,a,l\}} ||{\mathbf{f}^c_m}^T \mathbf{f}^u_m||_F^2 + \sum_{m \neq m' \in \{v,a,l\}} ||{\mathbf{f}^u_m}^T \mathbf{f}^u_{m'}||_F^2,$$

We formulate the data correlation loss:

$$\mathcal{L}_{enc_2} = \sum_{m \neq m' \in \{v,a,l\}} ||{\mathbf{f}^c_m}^T \mathbf{f}^c_{m'} - cI||_F^2,$$

The entire encoding loss:

$$\mathcal{L}_{enc} = \mathcal{L}_{enc_1} + \mathcal{L}_{enc_2}.$$

3. Decoding

The decoder is defined to ensure that the encoded representations indeed represent the details of the corresponding modality data.

$$\bar{\mathbf{f}}_m = D(\mathbf{f}^c_m + \mathbf{f}^u_m; \theta_d),$$

The reconstruction loss:

$$\mathcal{L}_{dec} = \sum_{m \in \{v,a,l\}} \frac{||\mathbf{f}_m - \bar{\mathbf{f}}_m||_2^2}{d_h},$$

Correlated Representation Learning (CRL)

The correlated representation learning can be achieved through the autoencoding architecture that is the combination of the encoders and the decoders.

$$\mathcal{L}_{CRL} = \alpha \mathcal{L}_{enc} + \beta \mathcal{L}_{dec},$$

4. Differential Privacy Protection Scheme

According to Basic Differential Privacy Mechanism, we can calculate the perturbed uncorrelated representation:

$$\hat{\mathbf{f}}_m^u = \mathbf{f}_m^u + Lap\left(0, S_{\mathbf{f}_m^u}/\epsilon\right),$$

According to Correlated Differential Privacy Mechanism, we can calculate the perturbed correlated representation:

$$\hat{\mathbf{f}}_m^c = \mathbf{f}_m^c + Lap\left(0, \sum_{m' \in \{v,a,l\}} Cos(\mathbf{f}_m^c, \mathbf{f}_{m'}^c) S_{\mathbf{f}_m^c}/\epsilon\right),$$

5. Privacy-Preserving Sentiment Prediction

We fuse the representation vectors into a joint vector, and then the prediction function is applied to the privacy-preserving prediction task:

$$\hat{\mathbf{y}} = G(\hat{\mathbf{f}}_{out}; \theta_{out}),$$

The Cross-Entropy Loss:

$$\mathcal{L}_{task} = -\frac{1}{n}\sum_{i=0}^{n} \mathbf{y}_i \cdot \log(\hat{\mathbf{y}}_i),$$

Differentially Private Correlated Representation Learning (DPCRL)

The overall loss of DPCRL:

$$\mathcal{L}_{DPCRL} = \alpha\mathcal{L}_{enc} + \beta\mathcal{L}_{dec} + \gamma\mathcal{L}_{task},$$

Experiments

Datasets: CMU-MOSI dataset and CMU-MOSEI dataset

Baselines: MISA, Self-MM, MMIM, and MISA-DP

Goal 1: (Compared with MISA, Self-MM, MMIM) To illustrate that our DPCRL model can maintain the sentiment analysis performance.

Goal 2: (Compared with MISA-DP) To illustrate that our DPCRL model outperforms the naive DP model.

Performance Metrics: Acc-2 and F1 (Neg/Non-neg), Acc-2 and F1 (Neg/Pos), and Acc-7

Evaluation Analysis: CRL Evaluation and DPCRL Evaluation

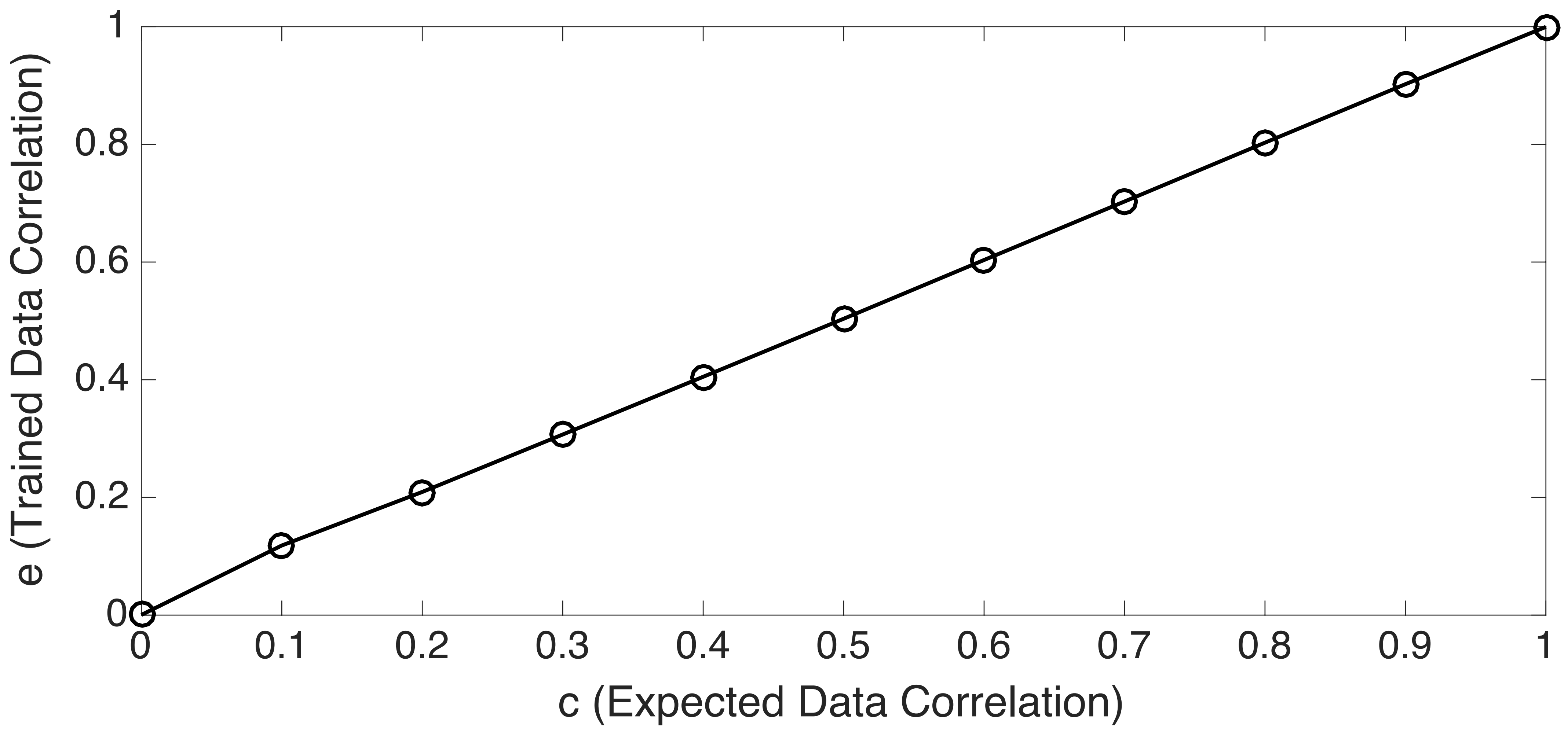
Goal 1: The impact of the expected correlation factor

Goal 2: The effectiveness of the proposed DPCRL

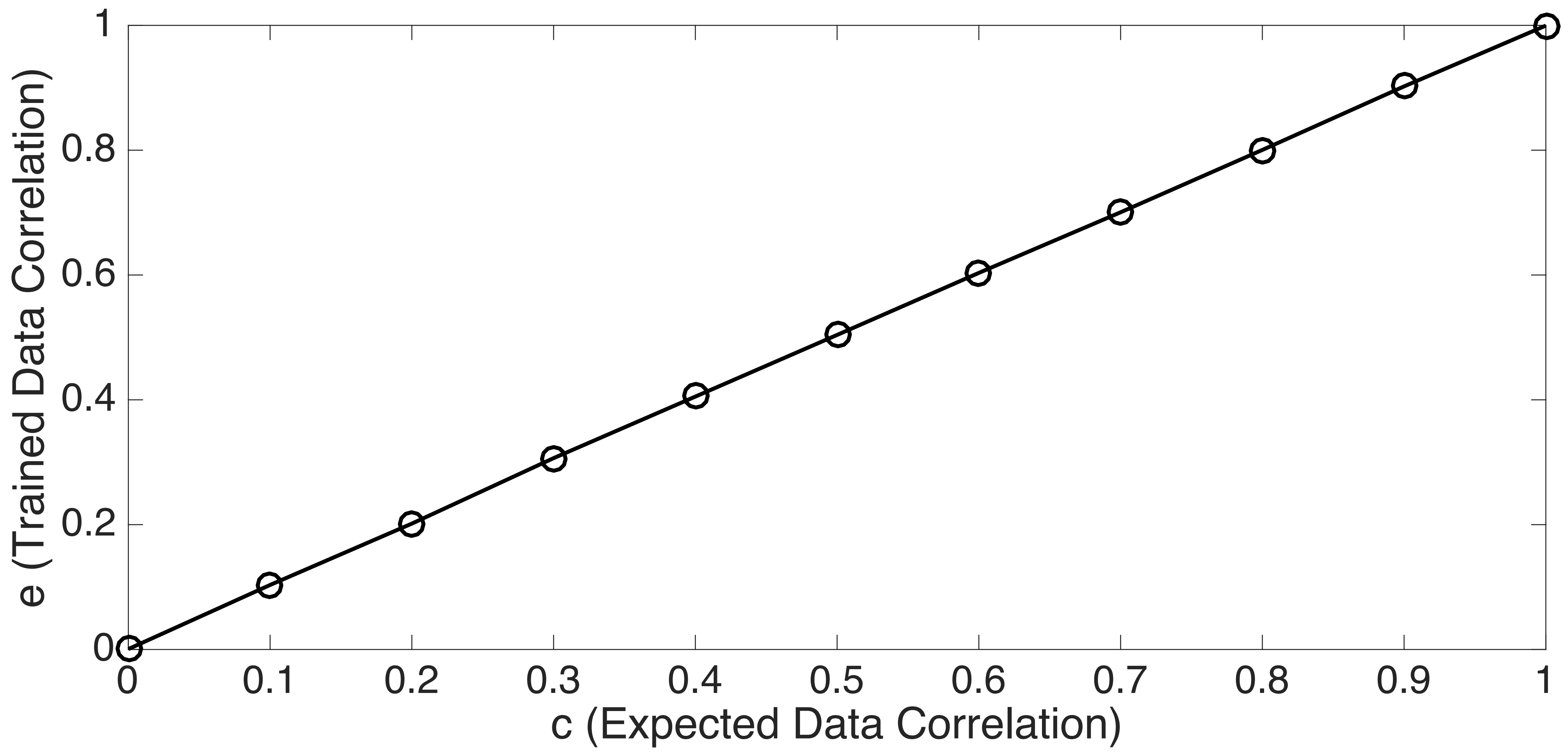
Evaluation on Correlated Representation Learning (CRL)

The impact of expected data correlation on trained data correlation:

MOSI Dataset



MOSEI Dataset

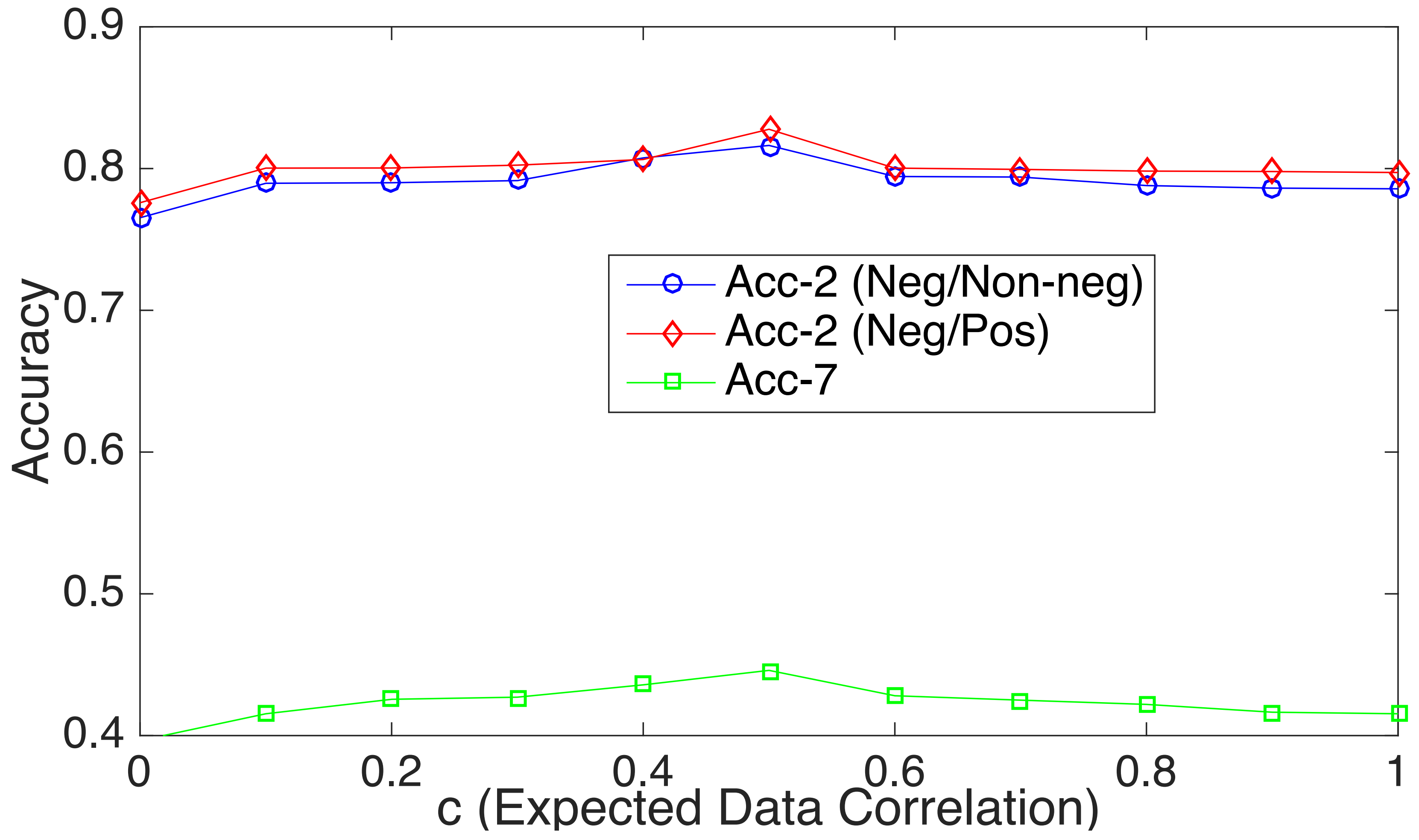


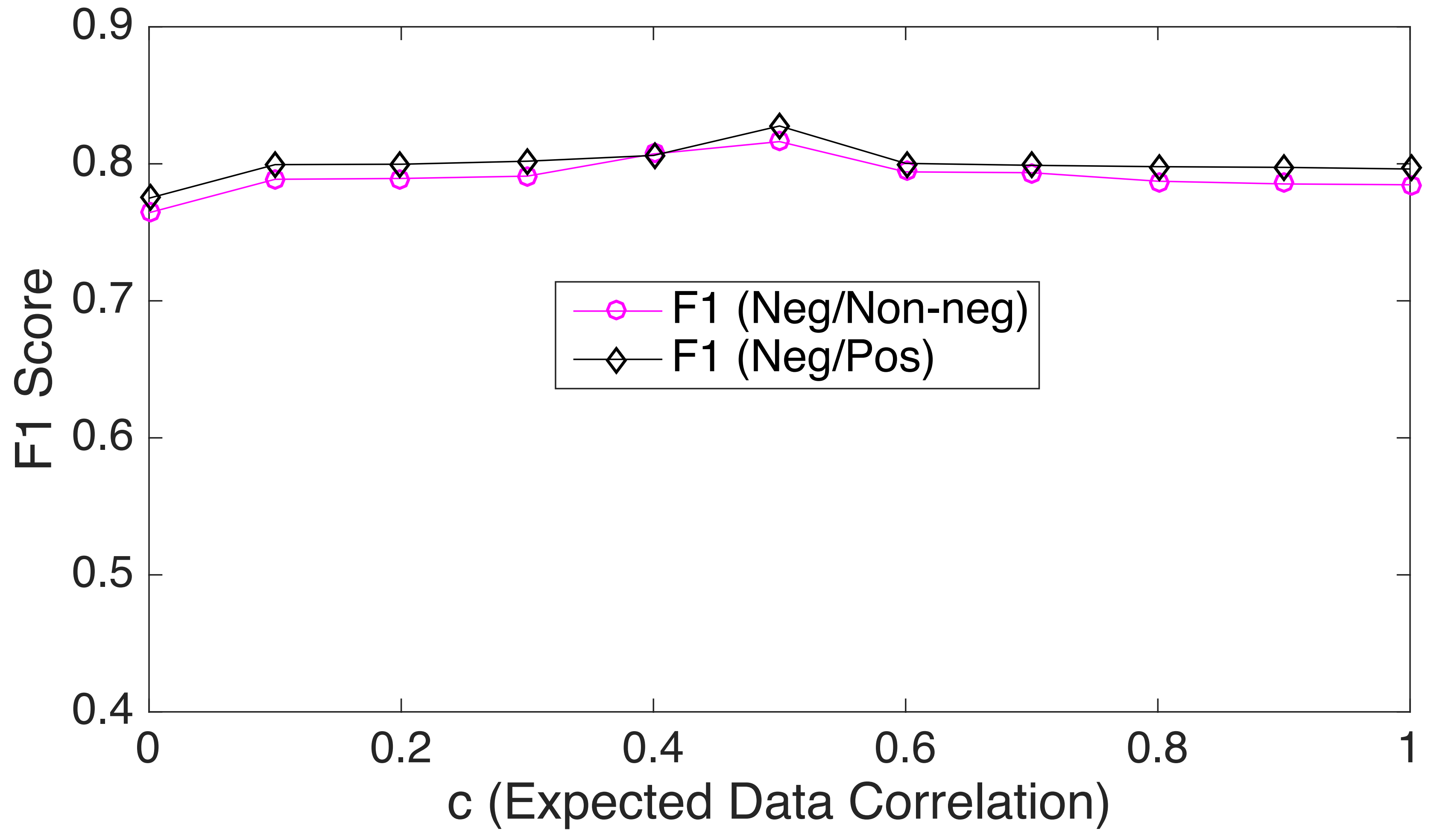
Remark 1: The results confirm that in our correlated representation learning scheme, the utilization of c is effective to accomplish our expected high-dimension data transformation.

Evaluation on Correlated Representation Learning (CRL)

The impact of expected data correlation on prediction accuracy of CRL:

MOSI Dataset

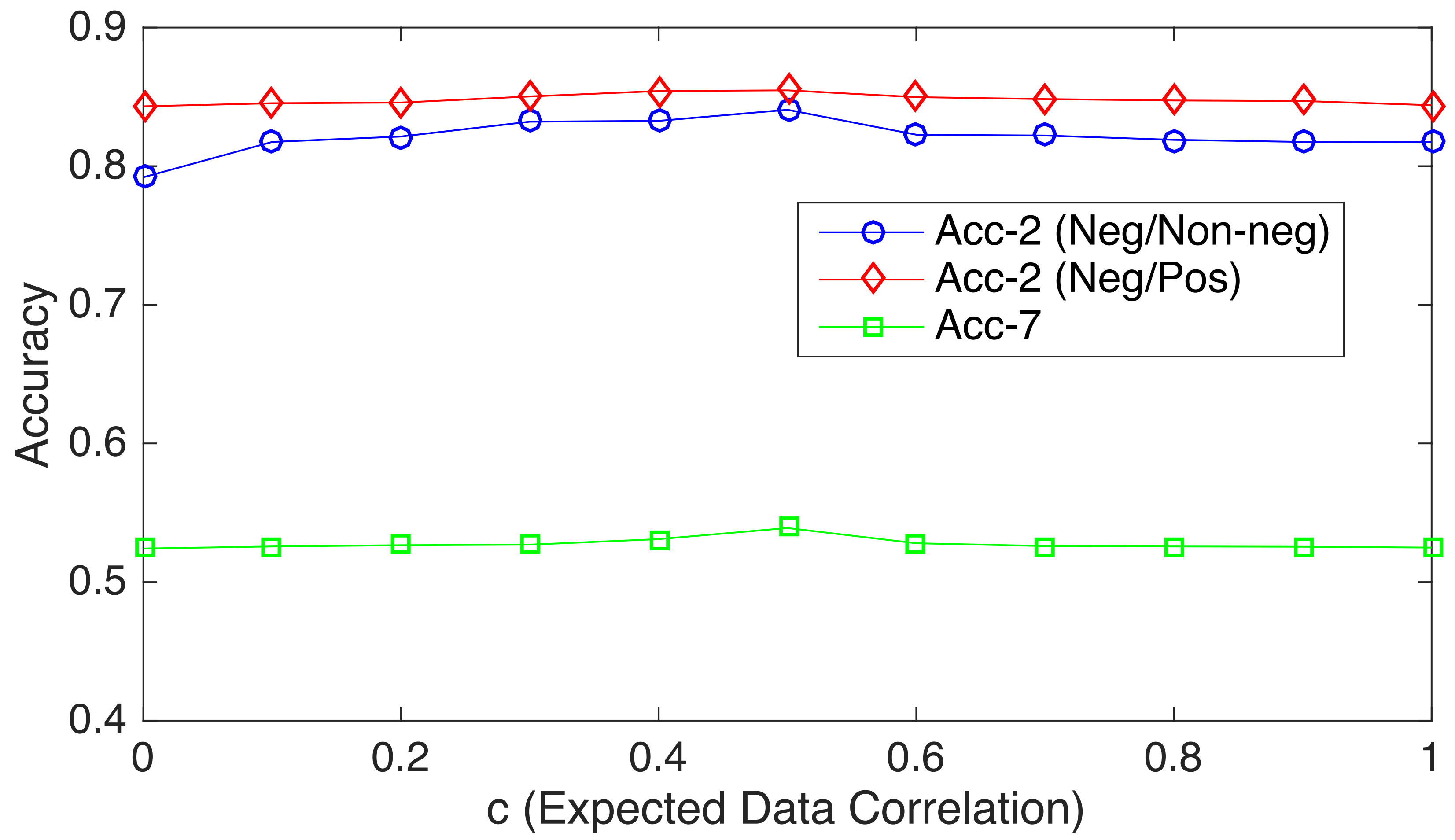


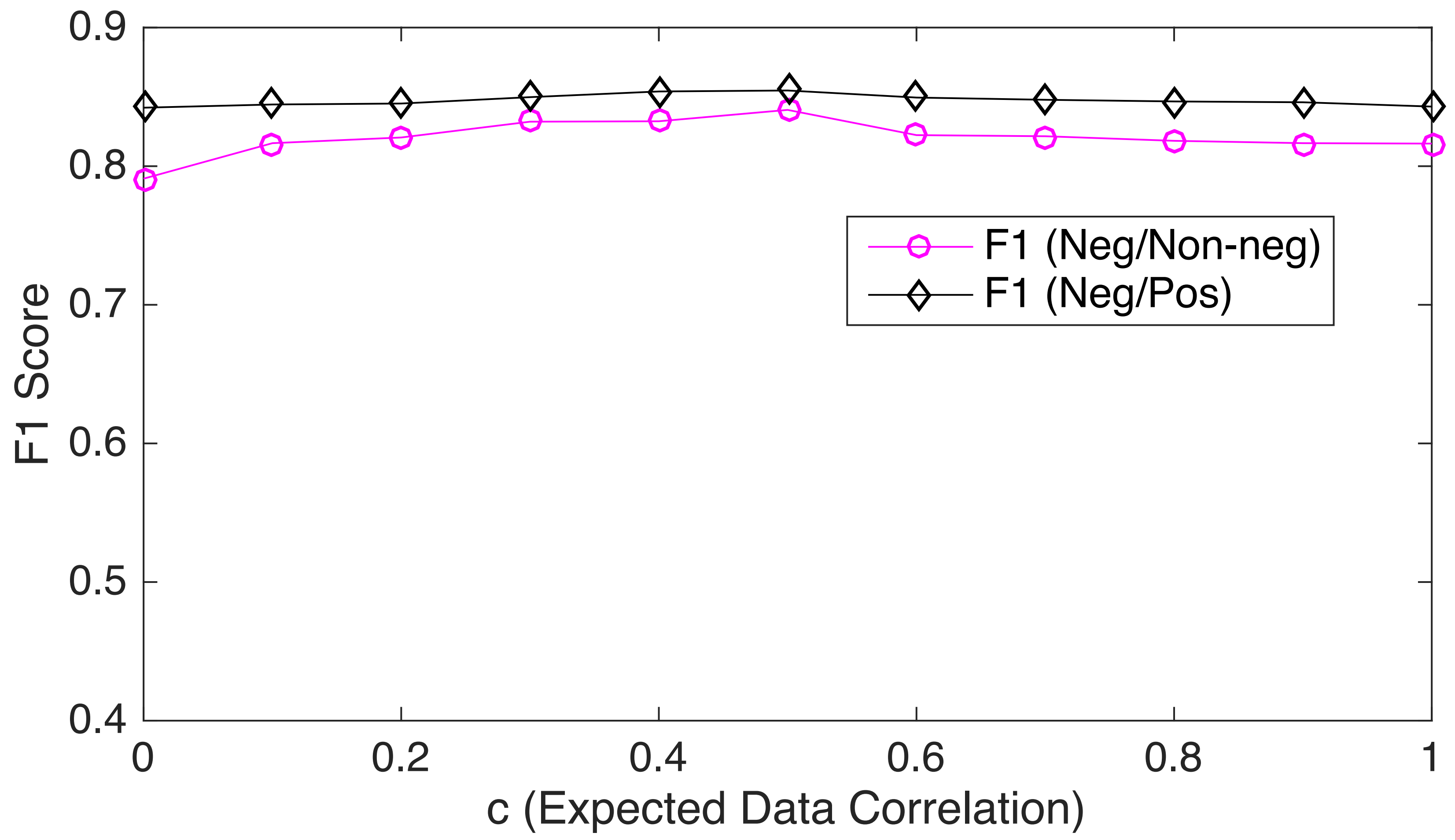


Evaluation on Correlated Representation Learning (CRL)

The impact of expected data correlation on prediction accuracy of CRL:

MOSEI Dataset



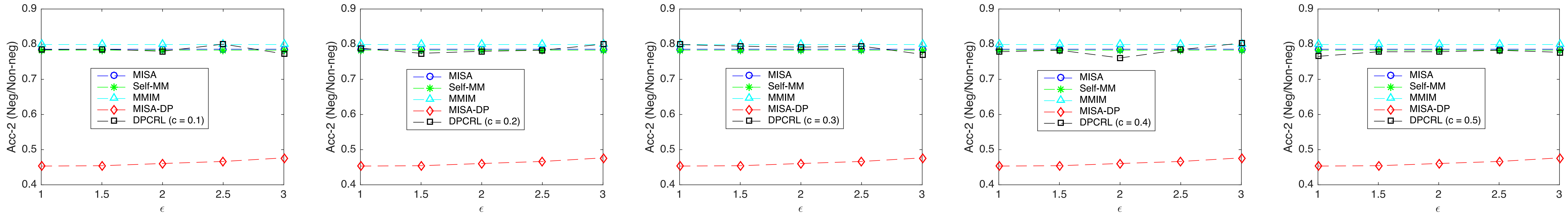


Remark 2: The correlation factor c can be used to balance the trade-off between representation similarity and representation diversity for improving multimodal sentiment analysis performance.

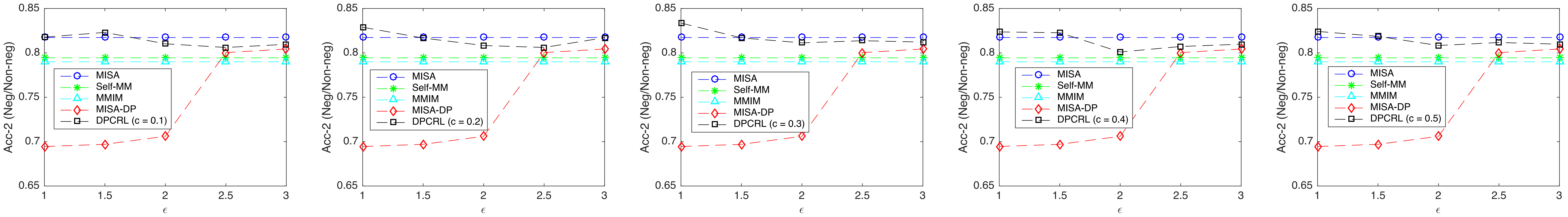
Evaluation on Our DPCRL Model

Georgia State University

Evaluation Results of Acc-2 (Neg/Non-neg) on MOSI Dataset (DPCRL vs. Baselines)

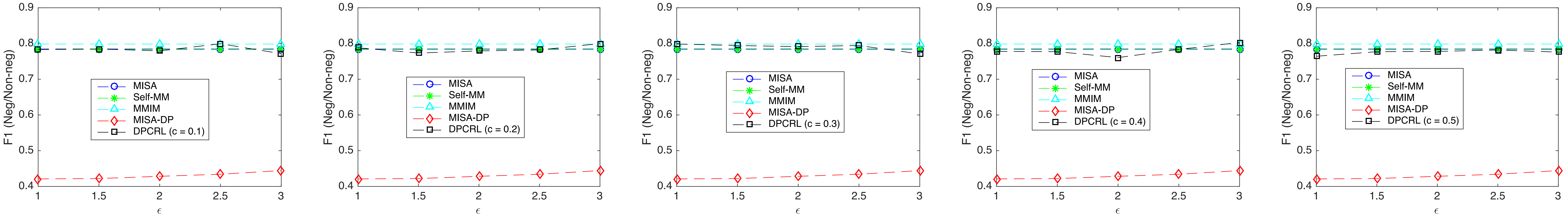


Evaluation Results of Acc-2 (Neg/Non-neg) on MOSEI Dataset (DPCRL vs. Baselines)

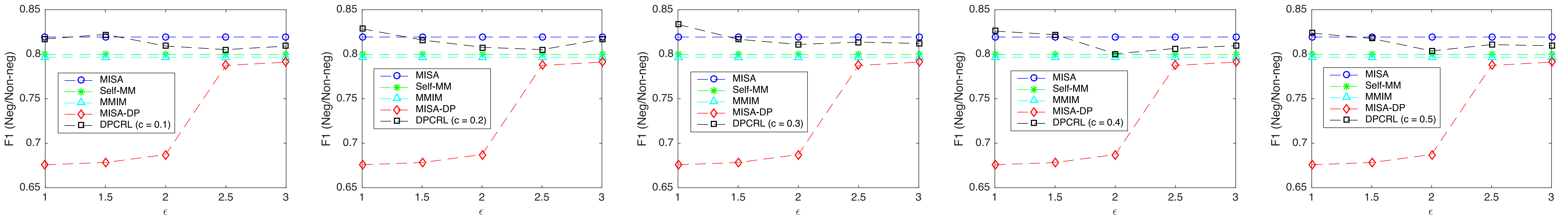


Evaluation on Our DPCRL Model

Evaluation Results of F1 (Neg/Non-neg) on MOSI Dataset (DPCRL vs. Baselines)

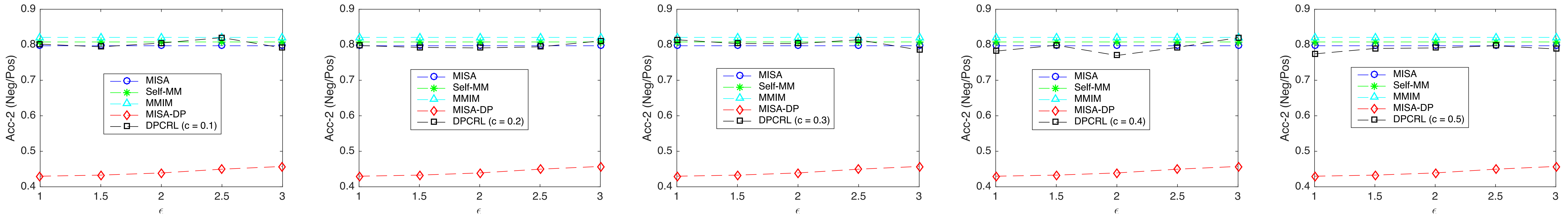


Evaluation Results of F1 (Neg/Non-neg) on MOSEI Dataset (DPCRL vs. Baselines)

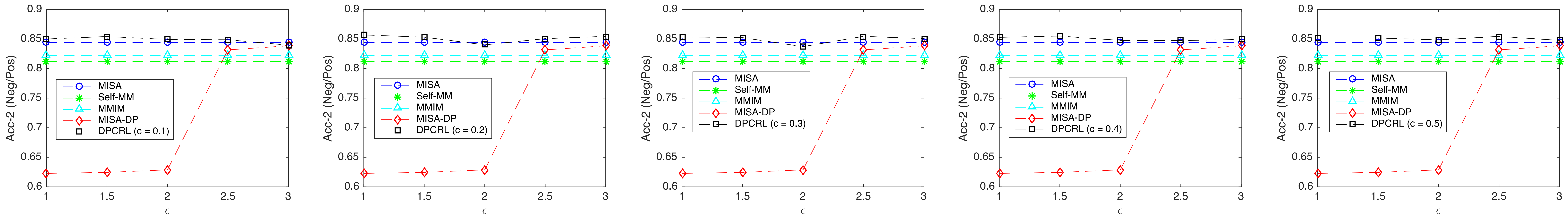


Evaluation on Our DPCRL Model

Evaluation Results of Acc-2 (Neg/Pos) on MOSI Dataset (DPCRL vs. Baselines)

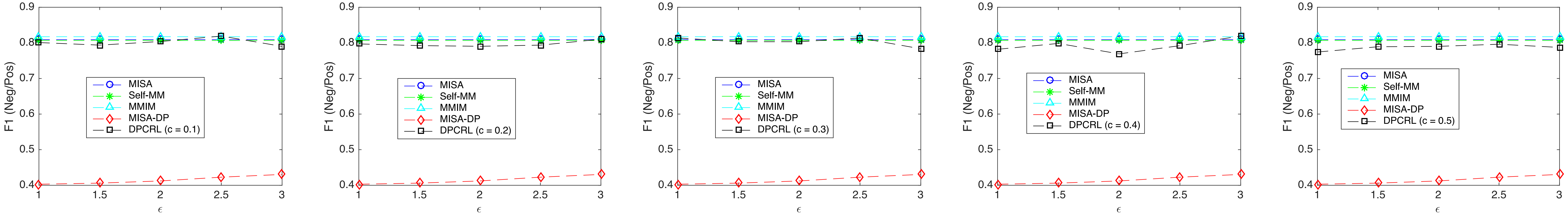


Evaluation Results of Acc-2 (Neg/Pos) on MOSEI Dataset (DPCRL vs. Baselines)

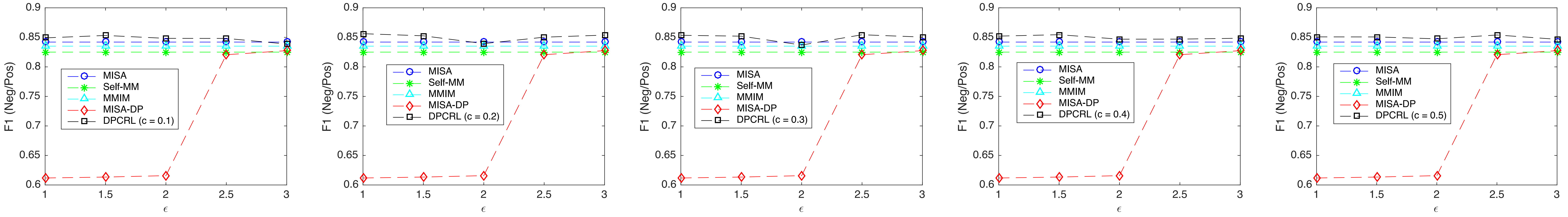


Evaluation on Our DPCRL Model

Evaluation Results of F1 (Neg/Pos) on MOSI Dataset (DPCRL vs. Baselines)

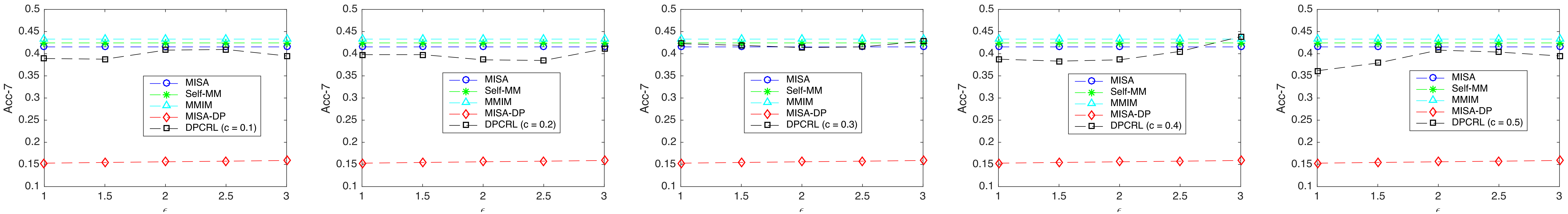


Evaluation Results of F1 (Neg/Pos) on MOSEI Dataset (DPCRL vs. Baselines)

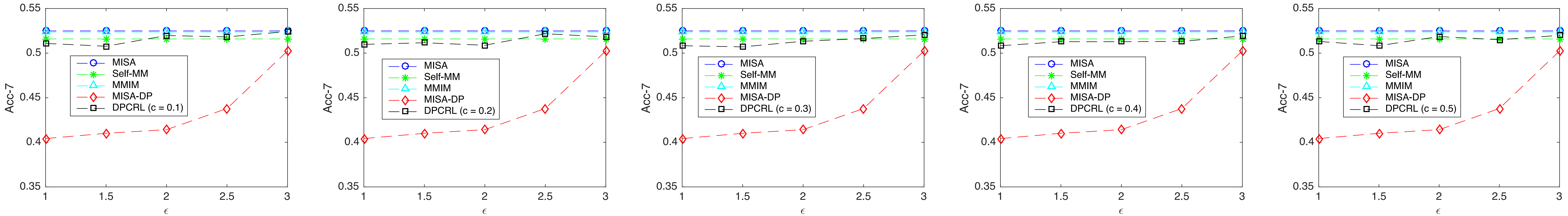


Evaluation on Our DPCRL Model

Evaluation Results of Acc-7 on MOSI Dataset (DPCRL vs. Baselines)



Evaluation Results of Acc-7 on MOSEI Dataset (DPCRL vs. Baselines)



Remark 3: DPCRL model can maintain the performance of sentiment analysis while satisfying differential privacy guarantee.

Remark 4: DPCRL can be leveraged to learn the correlated representations with a relatively lower correlation factor, mitigating the side-effect of the additional Laplace noise on the sentiment analysis.

Conclusion

1) This is the first work to design privacy-preserving multimodal sentiment analysis model.

2) Our proposed DPCRL model seamlessly combines a correlated representation learning scheme with a differential privacy protection scheme, aiming to simultaneously ensuring \epsilon-differential privacy and retaining the performance of multimodal sentiment analysis.

3) The high-dimension data transformation can be accomplished by learning the correlated and uncorrelated multimodal representations from multimodal data for sentiment prediction, and the expected correlation of correlated representations can be flexibly set via a correlation factor.

Thank you !!!