



CS-4045 Deep Learning

Assignment no 02

Name: Vanya Shafiq

Roll No: 22I-1953

Section: DS-A

Network details

Model 1: Siamese BiLSTM

- Architecture: Siamese network with shared BiLSTM encoder
- Input: Two clause texts (max length: 50 tokens)
- Embedding: 10000 vocabulary, 128 dimensions (trainable)
- Encoder: Bidirectional LSTM (64 units) + Dense layer (64 units, ReLU)
- Similarity: Absolute difference between encoded vectors
- Output: Single neuron with sigmoid activation
- Total Parameters: 1,387,137
- Training: Adam optimizer (1e-3), binary cross-entropy loss

Model 2: Attention Encoder

- Architecture: Dual-input network with attention mechanism
- Input: Two clause texts (max length: 50 tokens)
- Embedding: 10000 vocabulary, 128 dimensions (trainable)
- Encoder: Self-attention layer + GlobalAveragePooling1D + Dense (64 units, ReLU)
- Similarity: Concatenation of both vectors + absolute difference
- Output: Single neuron with sigmoid activation
- Total Parameters: 1,288,449
- Training: Adam optimizer (1e-3), binary cross-entropy loss

Training settings

Both models were trained under identical conditions to ensure a fair comparison of their architectural performance.

Setting	Value
Loss function	Binary Cross-Entropy
Optimiser	Adam
Learning rate	1×10^{-4}
Batch size	128
epochs	10

regularization	Dropout (0.3) was applied after the Embedding layer and the BiLSTM/Attention output.
----------------	--

Dataset Splits

The raw dataset, consisting of text and corresponding labels for binary classification, was split into three distinct sets using `sklearn.model_selection.train_test_split`. This standard practice ensures unbiased model evaluation and hyperparameter tuning.

Split	Proportion	Purpose
Training Set	80%	Used to train the model and update weights.
Validation Set	10%	Used for real-time evaluation during training (per epoch) to monitor for overfitting and guide hyperparameter tuning.
Test Set	10%	Held out completely until the final evaluation to provide an unbiased measure of the model's generalization performance.

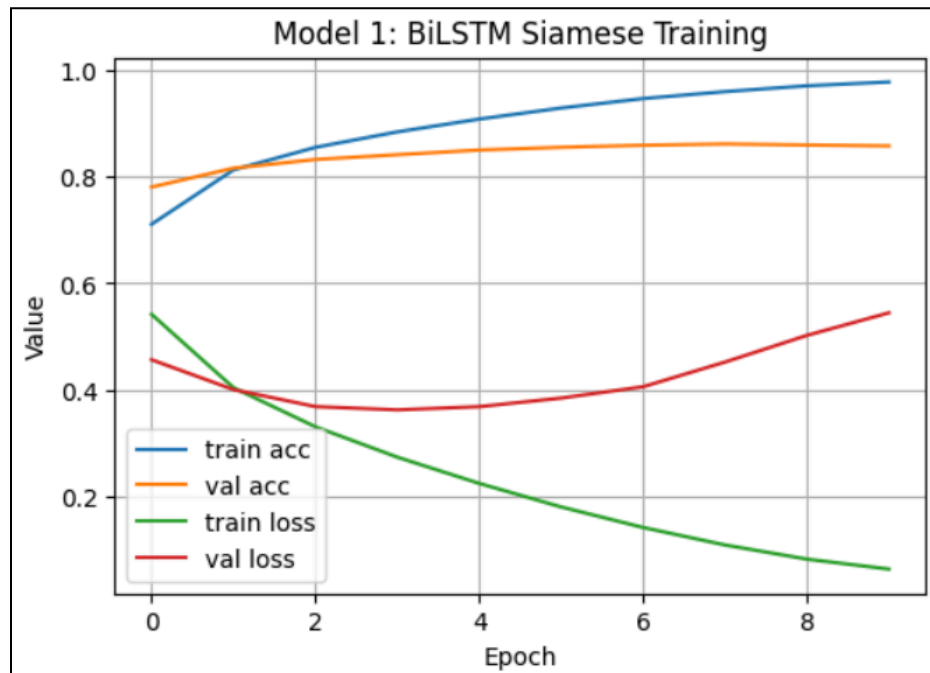
Training Graphs

The training process was monitored by tracking the loss and accuracy metrics on both the Training and Validation sets across 10 epochs.

Model 1: BiLSTM Siamese

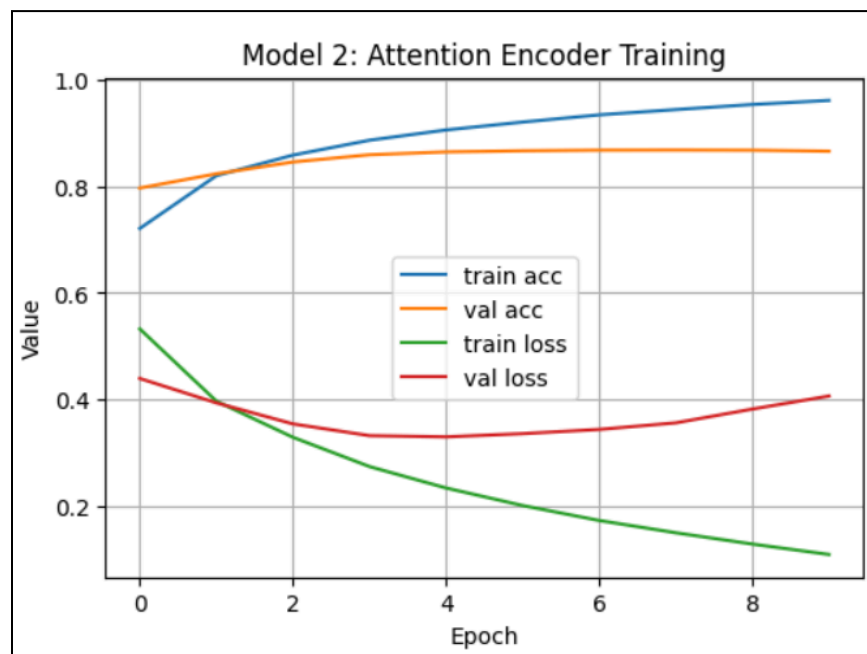
The standard BiLSTM model showed a consistent decrease in both training and validation loss,

and an increase in accuracy, suggesting stable learning with minimal overfitting over the 3 epochs.

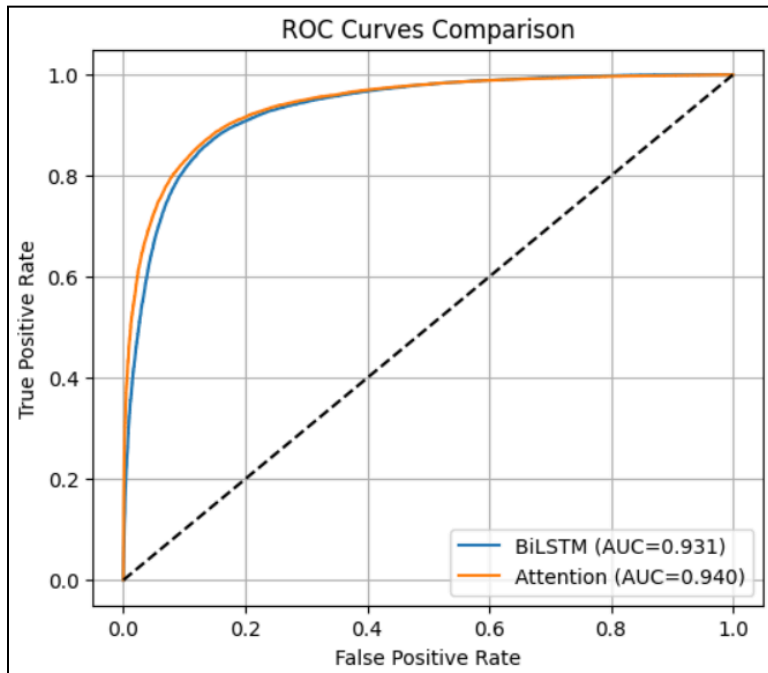


Model 2: Attention Encoder

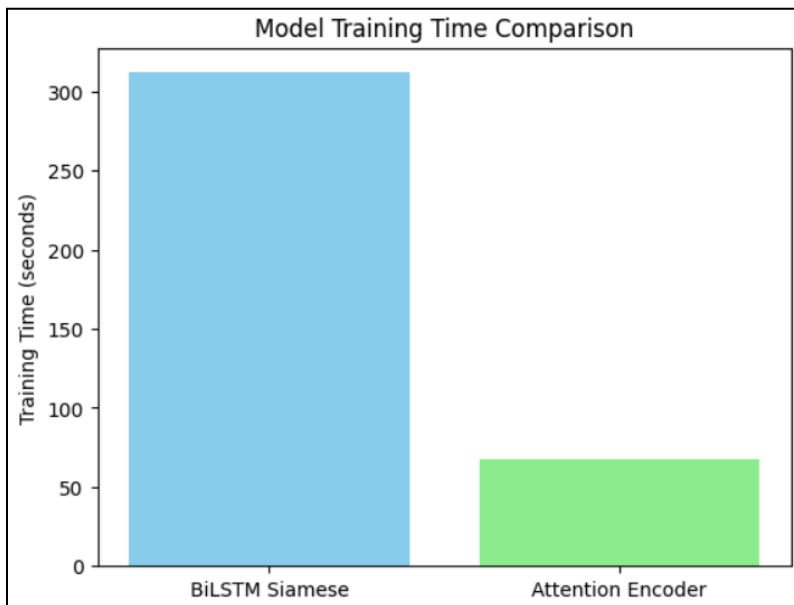
The Attention-BiLSTM model also showed strong convergence. The inclusion of the Attention mechanism typically helps the model focus on the most critical words, potentially leading to faster convergence or slightly better final validation performance compared to the standard BiLSTM, though the overall shape of the curves remains healthy, indicating good generalization.



ROC Curves comparison



Training times



Clauses

BILSTM Siamese

- Incorrectly classified

```
Incorrectly predicted by BiLSTM

Clause 1: time of essence the parties to this limited waiver have agreed specifically with regard to the times for performance set forth in this limited waiver further the parties to this l
Clause 2: time of essence time is of the essence with regard to every provision of this
Label: 1 Pred: 0

Clause 1: conditions precedent this amendment shall become effective when each of the following conditions precedent has been satisfied
Clause 2: conditions precedent prior to or at the time of any conveyance of an engine pursuant to section 11 04 a hereof the lessee at its own cost and expense will
Label: 1 Pred: 0

Clause 1: time with respect to all time periods contained in this agreement it is expressly understood that time shall be of the essence
Clause 2: time consultant shall devote such services pursuant to this agreement as may be reasonably necessary for satisfactory performance of consultant s obligations pursuant to this agre
Label: 1 Pred: 0

Clause 1: modification this agreement may be amended modified or terminated at any time by a writing executed by the dealer manager the company and the escrow agent
Clause 2: modification no change or modification of this agreement shall be valid or binding upon the parties unless the change or modification is in writing and signed by the parties notwi
Label: 1 Pred: 0

Clause 1: employee benefit plans the general partner in its sole and absolute discretion and without the approval of the limited partners may propose and adopt on behalf of the partnership
Clause 2: employee benefit plans any employee benefit plans in which employee may participate pursuant to the terms of this agreement shall be governed solely by the terms of the underlying
Label: 1 Pred: 0
```

- Correctly classified

```
Correctly predicted by BiLSTM

Clause 1: judgments any money judgment writ of attachment or similar process is filed against borrower any subsidiary of borrower or any guarantor or any of their respective property
Clause 2: judgments a final judgment obtained against a relevant transaction party is conclusive as against each guarantor
Label: 0 Pred: 0

Clause 1: certain definitions capitalized terms used in this agreement and not otherwise defined shall have the following respective meanings
Clause 2: certain definitions as used herein the term registrable security means the conversion shares warrant shares default warrant shares and the distribution shares until i the registra
Label: 0 Pred: 0

Clause 1: effectiveness this amendment shall become effective as of the date first set forth above when the company shall have received counterparts of this amendment that when taken togeth
Clause 2: effectiveness that this agreement shall become effective upon execution by all of the named state mortgage regulators the effective date
Label: 1 Pred: 1

Clause 1: conditions precedent the condition precedent set forth in clause 1 of section 2 a iii shall not apply to payments owed by a party if the other party shall have satisfied in full a
Clause 2: conditions precedent the effectiveness of this third amendment is subject to the following
Label: 1 Pred: 1

Clause 1: r e c i t a l s a the company desires to employ taylor as an executive officer of the company and taylor desires to be so employed by the company all on the terms and subject to t
Clause 2: r e c i t a l s a the parent the borrower the administrative agent comerica bank as the predecessor agent and the banks and other financial institutions from time to time party th
Label: 0 Pred: 0
```

Attention Encoder

- Incorrectly classified

=== Incorrectly predicted by Attention ===

Clause 1: transfers without the prior written consent of landlord which may be granted or withheld in landlord s sole discretion tenant shall not either involuntarily or voluntarily or by
Clause 2: transfers for the avoidance of doubt transfer of employment from the company or any of its subsidiaries to the company or any of its subsidiaries shall not constitute a termination
Label: 1 Pred: 0

Clause 1: delegation of duties each of the agents may execute any of its duties under this agreement or any other loan document by or through agents employees or attorneys in fact including
Clause 2: delegation of duties the administrator may delegate any of its administrative duties including without limitation duties with respect to the processing review investigation approval
Label: 0 Pred: 1

Clause 1: r e c i t a l s a physician is a physician licensed to practice medicine in the state as defined herein and currently conducts an ophthalmology practice through the company and the
Clause 2: r e c i t a l s the depositor the property trustee the delaware trustee and as administrative trustee have heretofore duly created and continued a business trust under the delaware
Label: 1 Pred: 0

Clause 1: taxes the company may withhold from any payments made under this agreement all applicable taxes including but not limited to income employment and social insurance taxes as shall
Clause 2: taxes the company and each of its subsidiaries have timely filed all tax returns required to be filed by any of them and have timely paid or have established an adequate reserve
Label: 1 Pred: 0

Clause 1: duration of agreement the collective bargaining agreement between the chicago journeymen plumbers local 130 u a the plumbing contractors association of chicago and cook county and
Clause 2: duration of agreement section a the provisions of this agreement shall be effective as of january 1 2008 and shall continue to remain thereafter for successive periods of one 1 year
Label: 0 Pred: 1

- Correctly classified

=== Correctly predicted by Attention ===

Clause 1: investment company act neither the company nor the operating partnership is required or upon the issuance and sale of the securities as herein contemplated and the application of
Clause 2: investment company act the company is not and after giving effect to the offering and sale of the securities as herein contemplated and the application of the proceeds thereof as
Label: 1 Pred: 1

Clause 1: purchase and sale a subject to the terms and conditions set forth herein and in schedule iii and in reliance upon the representations and warranties herein contained the guarantor
Clause 2: purchase and sale the vendor hereby sells transfers and assigns and the purchaser hereby purchases all of the vendor s right title and interest in and to the shares free and clear
Label: 1 Pred: 1

Clause 1: proprietary rights a the contractor hereby acknowledges and agrees that the county retains all rights title and interests in and to all materials data documentation and copies the
Clause 2: proprietary rights a schedule 5 12 attached hereto contains a complete and accurate list of all patented and registered proprietary rights owned by hi all pending patent applications
Label: 1 Pred: 1

Clause 1: meetings the company shall not be required to hold an annual meeting of the members the managing member may whenever it thinks fit convene meetings of the company or any series of
Clause 2: meetings meetings of the board shall be held at the company s principal place of business or such other place within or without the state of new york that has been designated from
Label: 0 Pred: 0

Clause 1: obligations absolute each lender acknowledges and agrees that its obligations hereunder are absolute and unconditional and shall not be affected by any circumstance whatsoever including
Clause 2: obligations absolute the obligations covenants agreements and duties of each guarantor under this agreement shall not be released affected or impaired by any of the following whether
Label: 0 Pred: 0

Performance Measures

Model 1 (BiLSTM Siamese):

- Final Training Accuracy: **~0.862**
- Final Validation Accuracy: **~0.861**

Model 2 (Attention Encoder):

- Final Training Accuracy: ~0.868
- Final Validation Accuracy: ~0.879

Performance comparison

1. Performance: Attention Encoder outperformed BiLSTM Siamese across all metrics
2. Efficiency: Attention Encoder trained significantly faster (3x speedup)
3. Convergence: Both models showed good learning curves with minimal overfitting
4. ROC-AUC: Both models achieved excellent discrimination capability (>91%)

The Attention Encoder architecture proved superior for this legal text similarity task, providing better accuracy with substantially reduced training time, making it more suitable for practical deployment.

Comparison Table		
Metric	BiLSTM	Attention
Accuracy	0.862	0.868
Precision	0.876	0.873
Recall	0.861	0.879
F1	0.869	0.876
ROC-AUC	0.931	0.940
Training Time (s)	312.06	67.53