

---

# TEXT CLASSIFICATION USING ATTENTION BASED BIDIRECTIONAL LSTM

---

A PREPRINT

✉ **Nabil Ananta Satria Gaharu**

2440035633

School of Computer Science  
Bina Nusantara University

nabil.gaharu@binus.ac.id

✉ **Devon Gasselyno**

2440026080

School of Computer Science  
Bina Nusantara University

devon.gasselyno@binus.ac.id

✉ **Jason Christian Wijaya**

2401963001

School of Computer Science  
Bina Nusantara University

jason.wijaya008@binus.ac.id

January 20, 2023

**Keywords** · LSTM · Text Classification

## 1 Introduction

Klasifikasi teks adalah proses pemberian kategori atau label yang telah ditentukan sebelumnya ke teks tertentu. Ini adalah kegiatan *supervised learning*, di mana model dilatih pada kumpulan data berlabel, yang mencakup sampel teks dan label yang sesuai. Tujuan model ini adalah mempelajari hubungan antara teks dan label, lalu menggunakan pengetahuan tersebut untuk mengklasifikasikan sampel teks baru yang belum pernah dilihat. (1)

*Spam detection* merupakan hal penting pada era digitalisasi, karena sejumlah besar pesan yang tidak diinginkan memenuhi *inbox* kita setiap hari. Metode tradisional untuk mendeteksi spam, seperti keyword-based filtering, menjadi semakin tidak efektif karena spammer ini terus mengembangkan teknik dan taktik mereka dalam beberapa tahun terakhir. (2)

Untuk mengatasi masalah ini, kami menggunakan model attention-based bidirectional LSTM untuk klasifikasi teks. Model ini menggabungkan mekanisme *attention*, yang memungkinkan network untuk menimbang pentingnya kata-kata yang berbeda dalam teks dan fokus pada fitur yang paling relevan untuk klasifikasi. Ini memungkinkan model untuk secara efektif menangkap semantik kompleks dan konteks teks, yang mengarah ke peningkatan kinerja. (3)

Dalam LSTM tradisional, jaringan memproses urutan input hanya dalam satu arah. Artinya, jaringan hanya dapat memperhitungkan konteks masa lalu saat membuat prediksi. Namun, dalam beberapa kasus, konteks masa depan juga bisa menjadi penting. Misalnya, dalam melakukan *Natural Language Processing*, kata berikutnya dalam sebuah kalimat dapat memberikan konteks penting untuk memahami arti kata saat ini. (4)(5)

Bidirectional LSTM (BLSTM) adalah jenis recurrent neural network (RNN) yang memproses urutan input dalam urutan maju dan mundur, memungkinkan model untuk memperhitungkan konteks masa lalu dan masa depan. Varian attention-based dari model BLSTM, menggunakan mekanisme *attention* untuk menimbang pentingnya berbagai bagian urutan input, dan berfokus pada bagian input yang paling relevan saat membuat prediksi. (6) Bidirectional LSTM dapat diterapkan dalam berbagai hal yang melibatkan data sekuensial, seperti *natural language processing*, *speech recognition*, dan *series forecasting*. (7)(8)

Mekanisme *attention* biasanya diimplementasikan sebagai komponen terpisah yang dilatih bersama BLSTM untuk mempelajari cara menetapkan bobot *attention* pada urutan input. Hal ini memungkinkan model untuk hadir ke bagian yang berbeda dari urutan input pada waktu yang berbeda, daripada memproses seluruh urutan dalam urutan yang sudah tetap. (9)(10)

Kami mengevaluasi model kami pada dataset dan akan membandingkan kinerjanya dengan *advanced* model lainnya. Hasil eksperimen menunjukkan keefektifan pendekatan yang kami usulkan dan menunjukkan bahwa pendekatan tersebut mengungguli metode lain dalam hal akurasi dan skor F1.

## 2 Previous Research

### 2.1 Literature Review

Title	Researcher	Method	Conclusion
Densely Connected Bidirectional LSTM with Applications to Sentence Classification	Zixiang Ding <sup>1</sup> , Rui Xia, Jianfei Yu, Xiang Li, Jian Yang <sup>1</sup>	Metode yang mereka gunakan dalam makalah ini adalah Densely Connected Bidirectional LSTM. Metode ini dibandingkan dengan Deep Stacked Bi-LSTM. perbedaan kedua model ini terletak pada bentuk dan hubungan antara layer masing masing RNN model. pada DC-Bi-LSTM setiap output layer RNN terhubung dengan layer RNN terakhir yang membuat proses pelatihan data lebih mudah.	Dense Connected Bidirectional LSTM adalah metode yang terdiri dari empat modul berupa input, dense Bi-LSTM, Average Pooling serta softmax layer. Berbeda dengan Deep Stacked Bi-LSTM, DC-Bi-LSTM lebih mudah untuk di train walaupun network yang luas dikarenakan setiap output layer RNN akan dikirim secara langsung pada layer RNN terakhir. selain itu model ini memiliki efisiensi parameter yang lebih baik dibandingkan Deep Stack Bi-LSTM atau model RNN pada umumnya dikarenakan untuk setiap layer RNN model ini dapat membaca input asli sequence sehingga tidak perlu memberikan semua informasi dan hanya perlu menambahkan informasi pada network. namun model ini juga memiliki kelemahan, karena hubungan dari (11)
A C-LSTM Neural Network for Text Classification	Chunting Zhou, Chonglin Sun, Zhiyuan Liu, Francis C.M. Lau	Metode yang mereka gunakan dalam makalah ini adalah metode C-LSTM (Buatan para researcher) yang merupakan gabungan dari convolutional neural network (CNN) dan long short term memory network (LSTM). Disini, CNN berguna mengekstrak urutan tingkat yang lebih tinggi dari fitur-fitur kata dan LSTM untuk menangkap ketergantungan jangka panjang atas urutan fitur masing-masing. Model ini mengubah setiap kalimat ke fitur berurutan (n-gram) untuk membantu menguraikan faktor variasi dalam kalimat.	C-LSTM dapat mempelajari fitur tingkat frasa melalui lapisan konvolusional; urutan representasi tingkat yang lebih tinggi tersebut kemudian dimasukkan ke dalam LSTM untuk mempelajari ketergantungan jangka panjang. Dinilai bahwa pembelajaran representasi kalimat semantik pada klasifikasi sentimen dan tugas klasifikasi tipe pertanyaan menghasilkan hasil yang sangat memuaskan. Hasil akhirnya sebanding dengan model-model yang sangat bergantung pada anotasi dan pengetahuan linguistik, terutama pohon parse sintaksis. Hal ini menunjukkan C-LSTM akan lebih layak untuk berbagai skenario. (12)

## 3 Methodology or Architecture Deep Learning

### 3.1 Model

Pada makalah ini, model yang kami gunakan akan didasari oleh model LSTM. Kami akan membandingkan akurasi dari tiga model lalu melakukan tuning pada code yang memiliki akurasi testing tertinggi. Berikut adalah Model yang akan kami gunakan :

Title	Researcher	Method	Conclusion
Attention-based Bidirectional LSTM for Deceptive Opinion Spam Classification	Ashish Salunkhe	Metode yang digunakan pada paper ini adalah Attention-based Bidirectional LSTM. Metode ini merupakan gabungan antara Bidirectional LSTM dengan metode attention-based. Perbedaan utama antara attention-based Bidirectional (BLSTM) dan metode LSTM lainnya adalah penggunaan mekanisme attention-based. sehingga berbeda dengan model LSTM lainnya, attention-based Bidirectional LSTM ini dapat untuk fokus pada bagian input yang paling penting sekaligus mempertimbangkan konteks masa lalu dan masa depan.	Attention-based bidirectional LSTM adalah arsitektur khusus yang menggabungkan mekanisme attention dengan bidirectional LSTM. Kombinasi ini memungkinkan model untuk fokus pada bagian input yang paling penting (menggunakan method attention based) sekaligus mempertimbangkan konteks masa lalu dan masa depan (menggunakan bidirectional LSTM). Dibandingkan dengan model lain, bidirectional LSTM perhatian memiliki beberapa keunggulan dalam mengklasifikasi spam yaitu, karena attention based memungkinkan model untuk fokus pada bagian masukan yang paling relevan, yang bisa sangat berguna saat menangani dokumen yang panjang dan rumit seperti email atau ulasan. Bidirectional LSTM memungkinkan model untuk mempertimbangkan konteks masa lalu dan masa depan, yang penting untuk memahami makna teks, terutama dalam kasus di mana makna kata tergantung pada konteksnya. Kombinasi perhatian dan bidirectional LSTM dapat meningkatkan kinerja dibandingkan dengan menggunakan salah satu metode saja. Namun, masih ada banyak model dan arsitektur lain yang dapat digunakan untuk klasifikasi spam seperti Random Forest, SVM, Naive Bayes, dll. Pemilihan model bergantung pada kasus penggunaan khusus dan data yang tersedia. Selain itu, penting juga untuk dicatat bahwa performa model apa pun bergantung pada kualitas dan kuantitas data pelatihan, serta detail implementasi spesifik. (13)

### 3.1.1 Bidirectional LSTM

Bidirectional LSTM adalah model yang didasari oleh LSTM. model ini sering digunakan untuk masalah sequential dikarenakan kemampuan model untuk memahami konteks sebuah kalimat dengan lebih akurat. Bidirectional LSTM memproses input dari 2 arah, kedua arah ini adalah arah depan dan belakang. Hasil dari output kedua arah akan digabungkan lalu dibuat untuk membuat prediksi. dengan begitu model ini dapat memeriksa konteks data yang lama dan baru saat memproses sebuah input. Hal ini dapat sangat membantu sebuah model untuk memahami subjek berupa *natural language processing* serta *speech recognition*. (14)(15)

### 3.2 Attention Based Bidirectional LSTM

Attention-based BLSTM cenderung mahal secara komputasi, terutama ketika berhadapan dengan urutan input yang besar. Mekanisme *attention-based* mungkin tidak selalu diperlukan, dan dalam beberapa kasus, model yang lebih sederhana dapat mencapai kinerja serupa dengan kompleksitas yang lebih sedikit. Model attention-based BLSTM dapat kesulitan jika tidak ada cukup data untuk melatih model. Model attention-based BLSTM sensitif terhadap pilihan mekanisme perhatian dan hyperparameter yang digunakan selama *training*, yang memerlukan tuning dan eksperimen yang teliti. Model attention-based BLSTM bisa jadi sulit untuk ditafsirkan dan dipahami dalam proses pengambilan keputusan dari model tersebut. (16)

Attention-based BLSTM memungkinkan model memperhitungkan konteks dari masa lalu dan masa depan, yang dapat berguna dalam task yang di mana konteksnya penting untuk memahami input. Mekanisme attention-based memungkinkan model untuk fokus pada bagian yang paling relevan dari urutan input, yang dapat sangat berguna

ketika berhadapan dengan urutan yang panjang, atau urutan di mana bagian tertentu lebih informatif daripada yang lain. Attention-based BLSTM dapat meningkatkan kinerja model pada *tasks* di mana urutan input memiliki *length* yang bervariasi dan model perlu memperhatikan bagian input tertentu. (17)

### 3.2.1 CNN LSTM

CNN-LSTM (Convolutional Neural Network-Long Short-Term Memory) adalah jenis arsitektur neural network yang menggabungkan kekuatan CNN dan LSTM. CNN-LSTM ini biasanya terdiri dari CNN yang digunakan untuk mengekstraksi fitur dari data input, diikuti oleh satu atau lebih lapisan LSTM yang memproses fitur yang diekstraksi dan menangkap dependensi temporal dalam data.(18)

CNN sangat bagus dalam mengekstrak fitur spasial dari gambar dan video, sementara LSTM bagus dalam menangkap dependensi temporal dalam data sekuensial seperti deret waktu, ucapan, dan teks. Dengan menggabungkan kedua jenis *network* ini, CNN-LSTM dapat memanfaatkan kekuatan CNN dan LSTM untuk menangani tugas yang melibatkan informasi spasial dan temporal. Lapisan CNN digunakan untuk mengekstrak fitur dari data input, seperti tepi, tekstur, dan bentuk. Fitur-fitur ini kemudian diteruskan ke lapisan LSTM, yang memproses fitur dan menangkap dependensi temporal dalam data. Keluaran dari lapisan LSTM kemudian dapat digunakan untuk tugas-tugas seperti klasifikasi atau prediksi. CNN-LSTM umumnya digunakan dalam tugas-tugas seperti keterangan gambar, klasifikasi video, dan pengenalan suara, di mana data input memiliki aspek spasial dan temporal.(19)(20)

Kelebihan utama menggunakan arsitektur CNN-LSTM (Convolutional Neural Network-Long Short-Term Memory) adalah menggabungkan kekuatan CNN dan LSTM untuk menangani tugas yang melibatkan informasi spasial dan temporal, CNN-LSTM juga dapat menangani berbagai skala informasi, seperti konteks global dan lokal, dan ketergantungan jangka pendek dan jangka panjang. Kelebihan lainnya juga termasuk arsitektur CNN-LSTM dapat menangani data yang kompleks seperti video, gambar, dan audio, dimana data tersebut memiliki aspek spasial dan temporal, hal ini dikarenakan lapisan CNN dapat mengekstraksi fitur spasial sedangkan lapisan LSTM dapat menangkap dependensi temporal, dan juga kemampuan CNN-LSTM untuk menangani dataset yang besar, *noise* dan *missing data*. (21)(22)

## 4 Analysis

### 4.1 Dataset

Untuk membandingkan hasil dari model kami, kami menggunakan dataset dari Yelp Business Reviews yang memberikan 12500 data ulasan tempat bisnis. data ini mengandung label serta ulasan dari tempat bisnis. label pada ulasan ini terbagi menjadi 2 kelas berupa 1 *truthful* atau 0 *deceitful*.

Untuk pelatihan dan percobaan, kami membagi 2 data dan mengambil data sebanyak 10000 untuk pelatihan dan 2500 untuk percobaan dengan rasio 80/20.

### 4.2 Preprocessing

Untuk Preprocessing kami melakukan pengecekan untuk *missing values* pada data, melabel kembali data sesuai kebutuhan, *stemming* kata pada data, menghilangkan *stopwords*, menghilangkan tanda baca dan normalisasi data. (12)

### 4.3 Model Training

Pada Model Bidirectional LSTM, kami menggunakan pretrained model GloVe(Global Vector for Word Representation) untuk melatih embedding layer yang nantinya akan dipakai pada model. selanjutnya kami menggunakan bidirectional layer yang didalamnya terdapat lstm layer. setelah itu diikuti oleh layer globalmaxpool dan dense layer untuk. untuk mengurangi *overfitting* kami menggunakan dropout layer dengan *drop out rate* sebesar 0.1. terakhir pada output layer, mengeluarkan 1 *unit* dengan adam optimizer dan binary crossentropy untuk loss function

Pada Model CNN-LSTM, model ini dimulai dengan mendefinisikan sebuah model *sequential* dengan *library* Keras, dan menambahkan beberapa *layer* ke dalam modelnya. *Layer* pertama yang ditambahkan ke model adalah sebuah 1D-Convolutional layer dengan output berdimensi *spatial* sama dengan inputnya, dilanjutkan dengan *dropout layer* untuk mencegah *overfitting*, dan menambahkan layer maxpooling yang digunakan untuk melakukan *down-sample* data input dengan mengambil nilai maksimal setiap 2 *timestep*. Selanjutnya model bidirectional LSTM ditambahkan, yang berguna untuk memproses data input dalam arah maju dan mundur dan menggabungkan output dari kedua arah. Ini memungkinkan model untuk menangkap konteks masa lalu dan masa depan dalam data input. Dense layer ditambahkan dengan softmax function untuk mengubah keluaran dari dense layer menjadi nilai probabilitas yang berjumlah 1.

Lapisan ini bertanggung jawab untuk klasifikasi akhir dari data masukan menjadi 4 kelas. Diakhiri dengan mengcompile model dan menentukan loss function, optimizer, dan metrics yang akan dipakai saat *training*.

Model Attention-based Bidirectional LSTM di *training* dengan metode *supervised learning*. Proses *training* melibatkan penyediaan model dengan dataset pasangan input output dan menyesuaikan parameter model untuk meminimalkan perbedaan antara output yang diprediksi dan output yang sebenarnya. (23) model ini terdiri dari beberapa *layer*, yang pertama adalah input *layer*, dilanjutkan oleh dua *embedding layer*, yang digunakan untuk mengubah data input menjadi representasi numerik. *Embedding layer* ini disertai juga dengan *dense layer* digunakan untuk menerapkan koneksi *dense* ke *concatenated output*. lalu terdapat Bi-LSTM layer di mana urutan input diproses oleh unit LSTM. Aspek bidirectional berarti ada dua lapisan LSTM, satu memproses urutan input ke depan dan satu memprosesnya ke belakang. Setelah itu kami juga menyediakan dua dropout layer yang digunakan untuk mencegah *overfitting* dengan menyetel sebagian kecil unit input secara acak ke 0 selama pelatihan. Kami juga menggunakan Batch Normalization untuk menormalkan aktivasi lapisan sebelumnya, yang dapat membantu menstabilkan proses *training*. Kemudian terdapat attention layer, di mana mekanisme attention diterapkan. Lapisan ini menghitung *weight* untuk setiap elemen dalam urutan masukan, menunjukkan betapa pentingnya bagi untuk prediksi saat ini, dan yang terakhir output layer di mana prediksi akhir dibuat berdasarkan output *weight* dari lapisan LSTM dan attention weight. (24) Setelah model sudah di *training*, model tersebut dapat digunakan untuk membuat prediksi pada urutan input baru. Mekanisme *attention* ini akan memungkinkan model untuk fokus pada bagian tertentu dari urutan masukan saat membuat prediksi, yang dapat membantu meningkatkan kinerjanya.

Secara keseluruhan kami tidak banyak merubah kode dari makalah referensi kami, kami mengubah dataset serta beberapa parameter dalam kode yang menurut kami dapat meningkatkan performa model.

#### 4.4 Result

Hasil dari model yang kami lakukan dianalisis dengan *performance metrics Precision, Recall, dan Accuracy*. Berikut adalah tabel perbandingan hasil model yang kami lakukan.

Accuracy		
Model	Train	Validation
Bi-LSTM	0.8674	0.8620
A-Bi-LSTM	<b>0.9983</b>	<b>0.8920</b>
Recurrent CNN	0.8590	0.8125

Loss		
Model	Train	Validation
Bi-LSTM	0.3056	0.3049
A-Bi-LSTM	<b>0.0119</b>	0.5631
Recurrent CNN	0.2441	<b>0.2688</b>

Berdasarkan hasil model didapatkan model Attention Based Bidirectional LSTM dengan akurasi train sebanyak 0.9983 dan *validation* sebesar 0.8920. A-Bi-LSTM juga memegang *train loss function* terkecil

#### 4.5 Hyperparameter Tuning

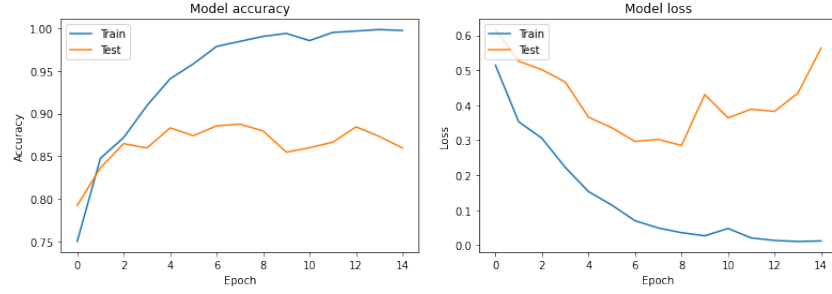
Hyper parameter tuning adalah sebuah proses untuk mencari kombinasi parameter model yang terbaik. Hyperparameter adalah parameter yang tidak ditemukan atau didapatkan dalam *model training* tetapi dibentuk sebelum *model training*. (25)

Berdasarkan hasil model pada result, didapatkan model Attention Based Bidirectional LSTM memiliki akurasi tertinggi serta *loss function* terendah. Sehingga untuk meningkatkan hasil model kami akan menggunakan *tuning* pada model tersebut. *Tuning* yang akan kami lakukan adalah Randomsearch tuning yang didapatkan dari *library* keras-tuner.

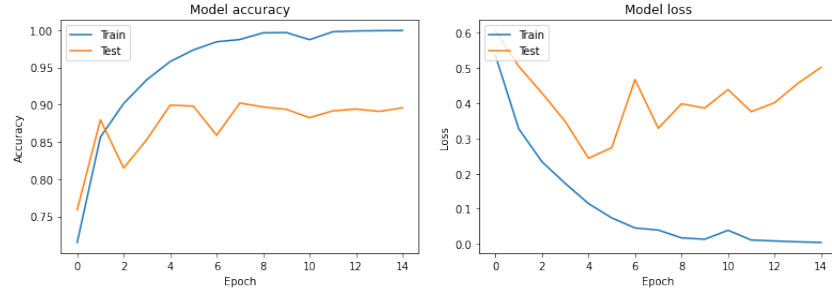
### 5 Evaluation

model Attention based Bidirectional LSTM

Attention-based Bidirectional LSTM Before Tuning



Attention-Based Bidirectional LSTM After Tuning



Loss		
Model	Train	Validation
Before	0.0119	0.5631
After	0.0043	0.5019

Accuracy		
Model	Train	Validation
A-Bi-LSTM	0.9983	0.8920
After	0.9993	0.8956

Berdasarkan gambar diatas dapat dilihat terdapat perubahan dan peningkatan dalam akurasi serta pengurangan dalam loss function model. Perubahan yang terlihat tidak begitu signifikan namun cukup terlihat ada peningkatan dari yang sebelumnya tidak dilakukan hypertuning. Untuk dapat lebih meningkatkan model ini dataset yang diberikan dapat diperbanyak untuk meningkatkan efisiensi serta ke akuratan prediksi dari model tersebut.

## References

- [1] M. Ikonakis, S. Kotsiantis, and V. Tampakas, "Text classification using machine learning techniques.," *WSEAS transactions on computers*, vol. 4, no. 8, pp. 966–974, 2005.
- [2] M. Crawford, T. M. Khoshgoftaar, J. D. Prusa, A. N. Richter, and H. Al Najada, "Survey of review spam detection using machine learning techniques," *Journal of Big Data*, vol. 2, no. 1, pp. 1–24, 2015.
- [3] Z. Huang, W. Xu, and K. Yu, "Bidirectional lstm-crf models for sequence tagging," *arXiv preprint arXiv:1508.01991*, 2015.
- [4] A. Graves, S. Fernández, and J. Schmidhuber, "Bidirectional lstm networks for improved phoneme classification and recognition," in *International conference on artificial neural networks*, pp. 799–804, Springer, 2005.
- [5] A. Kumar, V. T. Narapareddy, V. A. Srikanth, A. Malapati, and L. B. M. Neti, "Sarcasm detection using multi-head attention based bidirectional lstm," *Ieee Access*, vol. 8, pp. 6388–6397, 2020.
- [6] O. Melamud, J. Goldberger, and I. Dagan, "context2vec: Learning generic context embedding with bidirectional lstm," in *Proceedings of the 20th SIGNLL conference on computational natural language learning*, pp. 51–61, 2016.
- [7] A. Graves, N. Jaitly, and A.-r. Mohamed, "Hybrid speech recognition with deep bidirectional lstm," in *2013 IEEE workshop on automatic speech recognition and understanding*, pp. 273–278, IEEE, 2013.
- [8] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of lstm and bilstm in forecasting time series," in *2019 IEEE International Conference on Big Data (Big Data)*, pp. 3285–3292, IEEE, 2019.

- [9] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based lstm for aspect-level sentiment classification," in *Proceedings of the 2016 conference on empirical methods in natural language processing*, pp. 606–615, 2016.
- [10] M. Yang, W. Tu, J. Wang, F. Xu, and X. Chen, "Attention based lstm for target dependent sentiment classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, 2017.
- [11] Z. Ding, R. Xia, J. Yu, X. Li, and J. Yang, "Densely connected bidirectional lstm with applications to sentence classification," in *CCF International Conference on Natural Language Processing and Chinese Computing*, pp. 278–287, Springer, 2018.
- [12] C. Zhou, C. Sun, Z. Liu, and F. Lau, "A c-lstm neural network for text classification," *arXiv preprint arXiv:1511.08630*, 2015.
- [13] A. Salunkhe, "Attention-based bidirectional lstm for deceptive opinion spam classification," *arXiv preprint arXiv:2112.14789*, 2021.
- [14] Y. Liu, C. Sun, L. Lin, and X. Wang, "Learning natural language inference using bidirectional lstm model and inner-attention," *arXiv preprint arXiv:1605.09090*, 2016.
- [15] R. D. Deshmukh and A. Kiwelekar, "Deep learning techniques for part of speech tagging by natural language processing," in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 76–81, IEEE, 2020.
- [16] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "Abcdm: An attention-based bidirectional cnn-rnn deep model for sentiment analysis," *Future Generation Computer Systems*, vol. 115, pp. 279–294, 2021.
- [17] P. Zhou, W. Shi, J. Tian, Z. Qi, B. Li, H. Hao, and B. Xu, "Attention-based bidirectional long short-term memory networks for relation classification," in *Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers)*, pp. 207–212, 2016.
- [18] J. Wang, L.-C. Yu, K. R. Lai, and X. Zhang, "Dimensional sentiment analysis using a regional cnn-lstm model," in *Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers)*, pp. 225–230, 2016.
- [19] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using cnn-lstm neural networks," *Energy*, vol. 182, pp. 72–81, 2019.
- [20] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, "Action recognition in video sequences using deep bi-directional lstm with cnn features," *IEEE access*, vol. 6, pp. 1155–1166, 2017.
- [21] J. Zhang, Y. Li, J. Tian, and T. Li, "Lstm-cnn hybrid model for text classification," in *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 1675–1680, IEEE, 2018.
- [22] Q.-H. Vo, H.-T. Nguyen, B. Le, and M.-L. Nguyen, "Multi-channel lstm-cnn model for vietnamese sentiment analysis," in *2017 9th international conference on knowledge and systems engineering (KSE)*, pp. 24–29, IEEE, 2017.
- [23] J. Jiang, H. Zhang, C. Dai, Q. Zhao, H. Feng, Z. Ji, and I. Ganchev, "Enhancements of attention-based bidirectional lstm for hybrid automatic text summarization," *IEEE Access*, vol. 9, pp. 123660–123671, 2021.
- [24] Y. Bin, Y. Yang, F. Shen, N. Xie, H. T. Shen, and X. Li, "Describing video with attention-based bidirectional lstm," *IEEE transactions on cybernetics*, vol. 49, no. 7, pp. 2631–2641, 2018.
- [25] B. Wei, J. Li, A. Gupta, H. Umair, A. Vovor, and N. Durzynski, "Offensive language and hate speech detection with deep learning and transfer learning," *arXiv preprint arXiv:2108.03305*, 2021.