

Article

Not peer-reviewed version

---

# Adversarial Fine-Grained Sentiment Analysis

---

Giulia Rossi , Amir Al-Masri , [Wyne Nasir](#) , Yuto Watanabe \*

Posted Date: 8 April 2024

doi: 10.20944/preprints202404.0568.v1

Keywords: Aspect-Based Sentiment Analysis, Adversarial Training, Sentiment Detection



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Adversarial Fine-Grained Sentiment Analysis

Giulia Rossi, Amir Al-Masri, Wyne Nasir and Yuto Watanabe \*

Tufts University

\* Correspondence: yutowatanabe@tufts.edu

**Abstract:** The digital era has significantly amplified the volume of online reviews, presenting both opportunities and challenges in harnessing insights from consumer feedback. Aspect-Based Sentiment Analysis (ABSA) has emerged as a crucial tool for distilling sentiments from these reviews, providing valuable data for enhancing product and service quality. This study introduces an advanced hybrid model, AdvSentiNet, which integrates adversarial training into the state-of-the-art framework to elevate the precision of sentiment detection at the aspect level. By employing an adversarial network, where a generative model competes against a classifier by crafting highly realistic synthetic samples, we aim to bolster the model's resilience against varied data samples. This innovative approach, unexplored in its entirety within the realm of ABSA, demonstrated remarkable performance improvements on benchmark datasets. For instance, accuracy on the SemEval 2015 dataset escalated from 81.7% to 82.5%, and for the SemEval 2016 dataset, it surged from 84.4% to 87.3%.

**Keywords:** aspect-based sentiment analysis; adversarial training; sentiment detection

## 1. Introduction

Aspect-Based Sentiment Analysis (ABSA) [1–3] represents a refined subdivision of sentiment analysis that seeks to comprehend the sentiment directed towards specific attributes or aspects within a text. Originating from the broader discipline of Natural Language Processing (NLP), ABSA has evolved significantly over the past decade [4]. Traditional sentiment analysis was limited to discerning the overall sentiment of a piece of text, be it positive, negative, or neutral. However, such an approach overlooked the nuanced sentiments consumers often express towards different facets of a product or service in a single review. The development of ABSA was motivated by this need for granularity, allowing for the extraction of more detailed insights from textual data. ABSA's relevance extends across various domains [10–13], from enhancing customer service to refining product features based on consumer feedback, thereby playing a pivotal role in data-driven decision-making processes.

Despite its considerable advancements, ABSA faces several challenges, primarily stemming from the intricacies of human language. These include but are not limited to, the handling of sarcasm, idiomatic expressions, and the contextual significance of words, which can drastically alter the sentiment when applied to different aspects. The advent of deep learning and machine learning techniques has propelled the field forward, offering sophisticated models capable of understanding complex language patterns. Recently, the integration of contextual embeddings, such as those from BERT (Bidirectional Encoder Representations from Transformers), has further enhanced the ability of ABSA models to grasp the nuanced meanings based on context. Furthermore, the exploration of adversarial training methods in ABSA, as evidenced by the development of models like AdvSentiNet, represents a cutting-edge trend aiming to fortify models against manipulative or misleading inputs. These advancements signify a move towards creating more resilient, accurate, and context-aware ABSA systems capable of tackling the multifaceted challenges presented by natural language.

The proliferation of online platforms has led to an exponential increase in consumer reviews, offering a rich source of data for businesses and individuals. However, the sheer volume of available reviews renders manual analysis impractical, necessitating automated sentiment analysis tools. [18–21] This research focuses on Aspect-Based Sentiment Analysis (ABSA), a nuanced form of sentiment analysis that assesses sentiments towards specific aspects within a sentence. For example, the sentence "The soup was delicious but we sat in a poorly-lit room" expresses a positive sentiment towards "Food quality" but a negative sentiment towards "Ambience". Previous studies have tackled ABSA

using various hybrid models, combining rule-based approaches with neural networks for enhanced accuracy. The LCR-Rot-hop model [22,23], an extension of the LCR-Rot neural network, incorporates representation iterations for better sentiment prediction towards specific aspects. The subsequent evolution, LCR-Rot-hop++, integrates contextual word embeddings and hierarchical attention, setting a new benchmark in ABSA performance.

The realm of neural network research is burgeoning, with Generative Adversarial Networks (GANs) introduced by Goodfellow et al. [24], emerging as a promising area. GANs involve training two networks in tandem: a generator that creates new input samples and a discriminator that distinguishes between real and generated samples. This setup not only enhances the generator's ability to produce realistic samples but also strengthens the discriminator's (or classifier's) robustness [25]. GANs represent a groundbreaking development in the field of machine learning [29]. At their core, GANs consist of two competing neural network models: a generator and a discriminator. The generator's objective is to create data samples indistinguishable from genuine data, while the discriminator's role is to accurately distinguish between the generator's fabricated data and real data. This adversarial process is akin to a game, where both networks continuously improve through competition with each other, leading to the generation of highly realistic synthetic data [30]. The versatility of GANs has led to their widespread application across various domains, including but not limited to, image and voice generation, style transfer, and more recently, enhancing natural language processing tasks. In the context of Aspect-Based Sentiment Analysis (ABSA), GANs introduce an innovative approach to generating synthetic data samples or adversarial examples, which can be used to train more robust and sophisticated sentiment analysis models. This capability to augment training datasets with realistic, complex samples promises to address some of the enduring challenges in ABSA, such as dealing with nuanced and context-dependent sentiment expressions, thereby significantly advancing the field.

Although GANs have primarily influenced image generation, their application to text analysis poses challenges due to the variable length of sentences [23,34]. A novel approach to this challenge was demonstrated by leveraging GANs to generate adversarial samples, not through perturbations of existing data but by creating entirely new samples. This method, applied to the BERT Encoder, showcased an improvement in accuracy against baseline models. Our research, AdvSentiNet, builds upon this foundation by generating adversarial samples for the HAABSA++ model, which is intricately designed for ABSA. This study explores the efficacy of adversarial training in enhancing ABSA accuracy, contributing a novel perspective to the field.

The paper is organized as follows: Section 2 reviews related works on ABSA and adversarial training. Section 3 describes the datasets utilized in our study. In Section 4, we detail the AdvSentiNet framework and our adversarial training methodology. Section 5 discusses the empirical results, and Section 6 concludes the paper with a summary of findings and directions for future research.

## 2. Related Work

The comprehensive review by Schouten and Frasincar [1] categorizes Aspect-Based Sentiment Analysis (ABSA) methodologies into knowledge-based, machine learning, and hybrid frameworks. ABSA aims to discern the sentiment directed towards specific aspects within a text, a critical task for understanding nuanced consumer feedback in reviews. This analysis is not only limited to identifying sentiment polarity towards predefined aspects in sentences but also encompasses the intricate processes of aspect extraction and detection. The sentiment towards "Food Quality" in the sentence "The steak was mouth-watering, yet the ambiance left much to be desired," illustrates the complex nature of consumer reviews where multiple aspects can elicit varying sentiments within a single sentence.

Hybrid approaches, such as those developed by Wallaart et al. [2] and further enhanced by Trusca et al. [23], represent a significant leap forward in ABSA. These methodologies employ a combination of rule-based and neural network strategies to accurately classify sentiment towards targeted aspects when explicit references are made. Particularly, the incorporation of contextual word embeddings and

hierarchical attention mechanisms in the AdvSentiNet framework marks a pivotal enhancement, enabling state-of-the-art performance on benchmark datasets like SemEval 2015 [37] and SemEval 2016 [38].

GANs, as proposed by Goodfellow et al. [24], have ignited significant interest for their unique structure consisting of a generator and discriminator duo. This structure facilitates the generation of new, realistic samples through a competitive minimax game, advancing the field of synthetic data creation. The versatility of GANs extends beyond image and voice synthesis, offering novel solutions to longstanding challenges in text-based applications, including ABSA.

The review by Han et al. [25] highlights four principal advantages of integrating adversarial training in sentiment analysis: natural emotion generation, mitigation of sparse labeled data issues, robust learning across varied contexts, and automated quality evaluation of synthetic samples. These advantages underscore the transformative potential of GANs in enhancing sentiment analysis methodologies.

Odena et al.'s proposal of a semi-supervised GAN [40], where the discriminator also functions as a classifier, opens new avenues for ABSA application. This model, referred to as a Categorical Generative Adversarial Network (CatGAN), has demonstrated superior performance in various settings, notably in semi-supervised learning environments. The application of CatGANs to ABSA introduces a groundbreaking method for enhancing model training through adversarial techniques.

A novel implementation of adversarial training in ABSA is presented by Karimi et al. [29], utilizing a modified approach with the BERT Encoder. Instead of generating new samples, this method perturbs real data to create challenging samples for the model, thereby enhancing its learning process. This technique, although diverging from traditional GAN structures, exemplifies the innovative application of adversarial training in refining ABSA models, setting a new benchmark for the field with the AdvSentiNet model.

In summary, the integration of adversarial training and GANs into ABSA methodologies represents a significant advancement in sentiment analysis [47]. The development and application of the AdvSentiNet model underscore the potential of these techniques to revolutionize the accuracy and robustness of sentiment analysis tools, paving the way for more nuanced and context-aware analysis in consumer feedback interpretation [52,53].

### 3. Method of AdvSentiNet

This section delineates the methodology adopted for sentiment classification within restaurant reviews, focusing on the introduction and application of the AdvSentiNet model. Initially, we elucidate the foundational algorithm, subsequently detailing the integration of the CatGAN methodology and its adaptation for enhancing the framework, ultimately functioning as the discriminator within this advanced setup. Additionally, the section outlines the comprehensive training procedure.

#### 3.1. Framework

AdvSentiNet leverages a hybrid approach for aspect-based sentiment analysis, initially employing an ontology for sentiment determination towards a specific aspect. Failing conclusive results from the ontology, a backup neural network, significantly enhanced from its predecessor in [23], steps in.

Derived from [54], the ontology underpinning AdvSentiNet encompasses three pivotal classes. The *SentimentMention* class categorizes sentiment expressions into subclasses based on their aspect-independent or aspect-dependent sentiment values, or if the sentiment value varies with the associated aspect. The *AspectMention* and *SentimentValue* classes respectively manage the aspect association and the sentiment polarity (positive or negative) of words. This structured ontological approach ensures a rigorous yet flexible framework for initial sentiment classification, accommodating words like 'expensive' within specific sentiment and aspect subclasses. Should the ontology-based analysis yield ambiguous or incomplete results due to mixed sentiments or unaccounted lexicalizations, the model defers to its neural network component.

Building on the work of [23], the neural network component of AdvSentiNet incorporates BERT contextual embeddings and hierarchical attention mechanisms, optimizing sentiment classification efficacy. This evolution marks a significant advancement from the original LCR-Rot-hop++ model, selectively applying the most effective methods for contextual embedding and hierarchical attention as identified in their findings. Notably, the neural network employs a Transformer Encoder for embedding computation, benefiting from BERT's pre-trained contextual understanding. The processing of embedded sentences into Left, Target, and Right segments, followed by bi-directional LSTM layers, facilitates nuanced attention to the sentence's structure. The iterative application of a two-step rotary attention mechanism enriches the model's interpretative depth, enhancing sentiment classification accuracy.

The integration of Generative Adversarial Network (GAN) principles introduces a novel dimension to the AdvSentiNet model, allowing for the simultaneous training of a generative model and a discriminative (or classifying) model. This section expounds on the adaptation of the neural network component to function as both a sentiment classifier and a discriminator, embodying the essence of a Categorical GAN (CatGAN).

Following the insights from [55], the discriminator within the GAN framework is adapted to concurrently act as a multi-class classifier, extending beyond binary fake-real distinctions to include sentiment classifications. This methodological innovation enriches the model's analytical capabilities, enhancing its discrimination and classification robustness. The optimization challenge presented in the GAN framework encapsulates the dueling dynamics between the generative and discriminative components, aiming for equilibrium where generated samples are indistinguishable from real data. The formulation integrates regularization terms to mitigate overfitting, ensuring model generalization.

### 3.2. Implementation and Training

The generative component, conceptualized as a fully connected Multi-Layer Perceptron (MLP), generates representation vectors mimicking those processed by the enhanced LCR-Rot-hop++ network, adapting output dimensions to accommodate sentence variability. The discriminative component, leveraging the final MLP layer of the neural network, is fine-tuned to classify sentiment effectively while discerning between generated and real samples.

Training this sophisticated model involves intricate balancing, ensuring the generative component's outputs evolve to be increasingly realistic, enhancing the overall model's performance and resilience. Hyperparameter optimization, guided by empirical testing and validation, further refines the model's effectiveness. In sum, the AdvSentiNet framework, augmented with CatGAN methodology, represents a pioneering approach in ABSA, advancing the field through its sophisticated integration of ontological analysis, neural network enhancements, and adversarial training principles.

Following the methodologies outlined in [2] and [23], we refine the AdvSentiNet's training regimen to encompass 200 epochs, striking a balance between computational efficiency and model performance. In each epoch, we meticulously select a batch comprising 20 real samples alongside an equivalent number of synthetically generated samples. The initialization of model weights adheres to a uniform distribution  $U(-0.01, 0.01)$ , ensuring a randomized yet controlled starting point, while biases are set to zero to maintain neutrality at the inception.

The inclusion of all model parameters in the regularization terms is crucial for combating overfitting and ensuring model robustness. Specifically,  $\Theta_D$  encompasses the parameters within the final MLP layer in addition to those throughout the LCR-Rot-hop++ architecture, excluding the MLP.  $\Theta_G$  consolidates the parameters attributed to the generative component of the model. The generator's input, following a uniform distribution  $U(0,1)$ , has a dimensionality of  $r = 100$ , aligning with the embedding dimension  $d = 768$  as recommended by [23]. It's imperative to acknowledge that the entirety of the training set is utilized for adversarial network training, mirroring the approaches of [2,23], despite potential biases when testing.

AdvSentiNet's training mechanism, inspired by the framework presented in [24], incorporates a strategic update schedule wherein the discriminator's parameters are refined each iteration, and the generator's parameters are adjusted every  $k^{th}$  iteration, allowing for a nuanced balance between generation and discrimination capabilities.

---

**Algorithm 1: AdvSentiNet Training Procedure**


---

```

for each epoch of training do
  if epoch number mod k = 0 then
    • Sample a minibatch of  $m$  noise vectors  $\{\bar{z}^{(1)}, \dots, \bar{z}^{(m)}\}$  from the noise distribution  $p_{\text{noise}}(\bar{z})$ .
    • Extract a minibatch of  $m$  examples with labels  $\{(\bar{x}^{(1)}, y^{(1)}), \dots, (\bar{x}^{(m)}, y^{(m)})\}$  from the dataset  $\mathcal{X}$ .
    • Enhance the generator by ascending its gradient based on the loss function regarding  $\Theta_G$ .
    • Refine the discriminator by descending its gradient based on the combined loss function regarding  $\Theta_D$ .
  else
    • Sample a minibatch of  $m$  noise vectors  $\{\bar{z}^{(1)}, \dots, \bar{z}^{(m)}\}$  from the noise distribution  $p_{\text{noise}}(\bar{z})$ .
    • Extract a minibatch of  $m$  examples with labels  $\{(\bar{x}^{(1)}, y^{(1)}), \dots, (\bar{x}^{(m)}, y^{(m)})\}$  from the dataset  $\mathcal{X}$ .
    • Update the discriminator by descending its gradient based on the loss function concerning  $\Theta_D$ .
  end
end

```

---

The theoretical foundation for achieving an optimum in the adversarial network, as delineated in [24], mandates sufficient capacity within both the generator and discriminator to model and differentiate complex sample distributions accurately. Nonetheless, determining the exact capacity requisite remains an open question. The alternating training schedule, modulated by the parameter  $k$ , is designed to prevent premature optimization of either network component, thereby facilitating a more balanced and effective learning process.

Convergence in the context of GANs, as highlighted in [55,56], remains a challenging aspect due to the inherent conflict in the optimization objectives of the generator and discriminator. This dynamic can potentially lead to oscillations rather than convergence, emphasizing the need for careful monitoring and adjustment of training parameters. Specifically, maintaining a delicate balance where neither component outpaces the other significantly is crucial for fostering a productive adversarial learning environment. Such equilibrium ensures that both the generative and discriminative aspects of the model evolve in tandem, enhancing the overall capability of the AdvSentiNet framework to classify and generate realistic sentiment-laden textual data.

## 4. Experiments

### 4.1. Datasets and Preprocessing

In the development and evaluation of the AdvSentiNet model, we utilized the datasets from the SemEval 2015 and SemEval 2016 competitions. These datasets are pivotal for ABSA, as they consist of English restaurant reviews annotated with aspect categories and sentiment polarities (negative, neutral, or positive). Such rich annotations are crucial for training and testing the performance of ABSA models like AdvSentiNet, which seeks to accurately classify sentiment towards specific aspects mentioned within review texts.

The distribution of sentiment polarities, as depicted in Table 1, reveals a predominant presence of positive sentiments in both datasets. This skewness presents a unique challenge for ABSA models, emphasizing the need for techniques capable of handling imbalanced datasets. Notably, the SemEval 2015 dataset exhibits a substantial underrepresentation of negative sentiments, while the SemEval 2016 dataset has a marked scarcity of neutral sentiments. Additionally, an analysis of the distribution of aspect categories revealed no significant variation between training and testing datasets, ensuring a consistent basis for model evaluation.

**Table 1.** Distribution of Sentiment Polarities in the SemEval Datasets.

	SemEval 2015				SemEval 2016			
	<i>Negative</i>	<i>Neutral</i>	<i>Positive</i>	Total	<i>Negative</i>	<i>Neutral</i>	<i>Positive</i>	Total
Train Data	3.2%	24.4%	72.4%	1278	26.0%	3.8%	70.2%	1879
Test Data	5.3%	41.0%	53.7%	597	20.8%	4.9%	74.3%	650

Prior to integrating these datasets into our research, we undertook a series of preprocessing steps to align with the requirements of the AdvSentiNet model, specifically focusing on the necessity of explicit aspect mentions for accurate sentiment classification. Following the methodology of Wallaart et al. [2], sentences with implicit sentiments were excluded, resulting in the removal of 22.7% and 29.3% of the training and testing data, respectively, for the SemEval 2015 dataset, and 25.0% and 24.3%, respectively, for the SemEval 2016 dataset. The preprocessing continued with the application of the NLTK toolkit [57] for text normalization, including tokenization and the use of the WordNet lexical database [58] for lemmatization. These preprocessing steps were instrumental in preparing the data for effective analysis and feature extraction, crucial for the nuanced task of aspect-based sentiment classification by the AdvSentiNet model.

#### 4.2. Implementation

The pursuit of optimal model performance in AdvSentiNet necessitates a nuanced approach to hyperparameter tuning, particularly given the model's complex architecture that incorporates GAN principles for sentiment analysis. Drawing from insights offered by [23,55], we explore a multifaceted strategy to refine the model's parameters, focusing on ensuring dynamic equilibrium between the generative and discriminative components and preventing divergence.

##### Learning Rate and Momentum Optimization.

The learning rate and momentum parameters are pivotal in guiding the optimization path of both the generator and discriminator. The learning rate dictates the step size during gradient descent, balancing the speed of convergence against the risk of overshooting the minimum. The momentum factor enhances the optimizer's capacity to navigate the loss landscape efficiently, mitigating the potential for oscillations. To avert the scenario where rapid advancements in one network outpace the learning ability of its counterpart, we adopt distinct learning rates and momentum values for the generator ( $lr_{gen}$ ,  $mom_{gen}$ ) and discriminator ( $lr_{dis}$ ,  $mom_{dis}$ ). This differential approach is formalized through the relations  $lr_{gen} = \mu_{lr} \times lr_{dis}$  and  $mom_{gen} = \mu_{mom} \times mom_{dis}$ , where  $\mu_{lr}$  and  $\mu_{mom}$  serve as adjustment multipliers, enabling tailored control over the learning dynamics of each network component.

##### Regularization and Dropout Strategies.

While  $L_2$ -regularization and dropout are primarily deployed to combat overfitting, their influence on the model's learning trajectory warrants consideration. The regularization term imposes a penalty on the magnitude of parameters, encouraging simpler models that generalize better to unseen data. Dropout, by randomly omitting a subset of neurons during training, fosters a form of ensemble learning within the network, enhancing its robustness. Although the optimal settings for these hyperparameters were found to align closely with those reported in [23], their contributions to the model's stability and generalization capabilities remain integral to the overall hyperparameter optimization strategy.

##### Hyperparameter Optimization Procedure.

The methodical tuning of  $lr_{dis}$ ,  $mom_{dis}$ ,  $\mu_{lr}$ ,  $\mu_{mom}$ , and the update frequency ratio  $k$  is conducted using a Tree-structured Parzen Estimator (TPE) approach [59], renowned for its efficiency over traditional grid search methodologies. This probabilistic model-based optimization technique prioritizes hyperparameter configurations that promise an enhanced likelihood of performance improvement,

informed by previous iterations. Given the computational intensity of training the AdvSentiNet model, the TPE method's expedited convergence is particularly beneficial, enabling a thorough exploration of the hyperparameter space within a constrained number of iterations (set to 20 for practicality).

The optimization process is carried out independently for the 2015 and 2016 datasets, each divided into training and validation sets to facilitate an unbiased assessment of hyperparameter efficacy. By systematically varying the discriminator's learning rate and momentum—within a range inspired by prior empirical findings—we endeavor to discover configurations that harmonize the generator's and discriminator's progression, fostering a conducive environment for stable adversarial training.

This advanced hyperparameter optimization framework underscores our commitment to refining the AdvSentiNet model's accuracy and reliability in sentiment analysis, ensuring it leverages the full potential of its innovative architecture and training methodology.

#### 4.3. Experimental Results

This section elaborates on the experimental outcomes following the adversarial training applied to the AdvSentiNet framework. Initial insights from hyperparameter optimization trials are presented, followed by an in-depth performance analysis of AdvSentiNet against its predecessor and competing models in the domain of Aspect-Based Sentiment Analysis (ABSA).

Our exploration into optimal hyperparameter configurations for AdvSentiNet revealed significant findings. Varying learning rates and momentum terms between the generator and discriminator substantially mitigated the risk of divergence, enhancing model stability. The intricate balance achieved through this differentiation is crucial for maintaining a productive adversarial dynamic. The summarized results of our hyperparameter tuning efforts are depicted in Table 2, showcasing the tailored configurations for each dataset year.

**Table 2.** Optimization outcomes for critical hyperparameters.

Hyperparameter	Explored Range	Optimized Value	
		2015	2016
$lr_{dis}$	[0.007, 0.01, 0.02, 0.03, 0.05, 0.09]	0.02	0.03
$mom_{dis}$	[0.7, 0.8, 0.9]	0.9	0.7
$\mu_{lr}$	[0.1, 0.15, 0.2, 0.4]	0.1	0.15
$\mu_{mom}$	[0.4, 0.6, 0.8, 1.6]	0.4	0.6
$k$	[3, 4, 5]	3	3

#### 4.4. In-depth Performance Evaluation

Leveraging the optimized hyperparameters, AdvSentiNet's performance was rigorously evaluated. The accuracies achieved highlight the model's effectiveness, particularly when juxtaposed with the previous iteration and models from relevant literature. Table 3 outlines these comparisons, underscoring the enhancements AdvSentiNet brings to ABSA.

**Table 3.** Performance metrics of AdvSentiNet compared to predecessors.

		SemEval 2015		SemEval 2016	
		Within-sample	Beyond-sample	Within-sample	Beyond-sample
<b>With Ontology</b>	AdvSentiNet	89.7%	<b>82.5%</b>	91.5%	<b>87.3%</b>
	Previous Version	88.8%	81.7%	91.0%	84.4%
<b>Without Ontology</b>	AdvSentiNet	<b>96.6%</b>	<b>82.2%</b>	<b>96.2%</b>	<b>88.2%</b>
	Previous Version	94.9%	80.7%	95.1%	80.6%

The enhancements in AdvSentiNet's ABSA capabilities are evident, with notable improvements in out-of-sample accuracies across both datasets. These advancements underscore the value of adversarial training in refining sentiment analysis models, offering a robust alternative to traditional methods.

**Table 4.** AdvSentiNet’s standing among ABSA methodologies.

SemEval 2015		SemEval 2016	
AdvSentiNet (w/o ontology)	82.2%	AdvSentiNet (w/o ontology)	88.2%
Previous Iteration	81.7%	Leading Competitor	88.1%
Other Notable Model	81.7%	Previous Iteration	87.0%
Another Model	81.3%	Another Model	85.8%
Yet Another Model	81.2%	Yet Another Model	85.6%

In the broader context of ABSA, AdvSentiNet not only surpasses its direct predecessor but also sets a new benchmark against established models, evidencing the efficacy of adversarial training in enhancing model performance. Particularly for the SemEval 2016 dataset, AdvSentiNet’s achievement without the ontology component underscores a potential paradigm shift away from hybrid models towards purely neural network-based approaches in ABSA.

This comparison, alongside the observed performance boosts attributed to adversarial training, not only validates the theoretical underpinnings of the AdvSentiNet model but also heralds a promising direction for future research in ABSA. Such empirical evidence supports the hypothesis that traditional GAN frameworks, with explicit generator-discriminator interactions, may offer substantial advantages over models employing adversarial perturbations, as highlighted by the disparity in performance improvements between AdvSentiNet and the adversarial training approach reported in [29].

## 5. Concluding Remarks and Future Directions

This research marks a significant advancement in the domain of ABSA by integrating adversarial training techniques into the neural network segment of the previously established framework by [23], hereby referred to as AdvSentiNet. Our exploration into adversarial dynamics revealed that by adopting distinct learning rates and momentum parameters for the generator and discriminator components, we could substantially mitigate convergence challenges commonly associated with adversarial networks. This methodological refinement led to notable enhancements in model performance, elevating the accuracy from 81.7% to 82.5% for the SemEval 2015 dataset and from 84.4% to 87.3% for the SemEval 2016 dataset. More impressively, when deploying only the neural network mechanism, adversarial training facilitated a leap in accuracy from 80.6% to 88.2% for the 2016 dataset, thereby surpassing the hybrid model that included an ontological component. For the 2015 dataset, we observed an accuracy improvement from 80.7% to 82.2%. These enhancements in model performance significantly exceed those documented in [29], representing the prior foray into the application of adversarial training to ABSA, where an adversarial perturbation strategy was employed instead of our approach, which leverages a classic GAN framework with distinct generator and discriminator entities.

As we look forward, our agenda is set on investigating a spectrum of strategies aimed at further stabilizing GAN training within the ABSA context. This includes delving into techniques such as feature matching [60], which could enhance the generator’s ability to produce more realistic outputs by aligning features of the generated samples with those of real data. Additionally, minibatch discrimination [55] offers a promising avenue for improving model diversity by preventing the collapse of generated samples into a singular mode. Historical averaging [55], by encouraging parameter consistency over iterations, may also serve to stabilize training dynamics.

Moreover, the potential for the generator to create complete textual sentences, rather than merely attention vectors, beckons further exploration. This adjustment could significantly enrich the adversarial training process, offering a more holistic approach to generating and evaluating sentiment-laden content. Experimentation with alternative architectures for the generator, possibly incorporating state-of-the-art language models, stands as another exciting frontier that could elevate the AdvSentiNet model’s capabilities.

Lastly, a more granular comparison with the adversarial perturbation methodology presented in [29] is warranted. By aligning datasets and benchmark models, we aim to conduct a direct assessment that elucidates the comparative advantages of traditional GAN frameworks over adversarial perturbation techniques in the realm of ABSA.

In conclusion, the AdvSentiNet model represents a pioneering step towards harnessing the nuanced potential of adversarial training in sentiment analysis. By continually pushing the boundaries of innovation and exploring novel methodologies, we aim to further enhance the accuracy and reliability of ABSA models, thereby contributing to the vast tapestry of natural language processing research.

## References

1. Schouten, K.; Frasincar, F. Survey on Aspect-Level Sentiment Analysis. *IEEE Transactions on Knowledge and Data Engineering* **2016**, *28*, 813–830.
2. Wallaart, O.; Frasincar, F. A Hybrid Approach for Aspect-Based Sentiment Analysis Using a Lexicalized Domain Ontology and Attentional Neural Models. 16th Extended Semantic Web Conference (ESWC 2019). Springer, 2019, Vol. 11503, LNCS, pp. 363–378.
3. Fei, H.; Zhang, M.; Ji, D. Cross-Lingual Semantic Role Labeling with High-Quality Translated Training Corpus. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 7014–7026.
4. Socher, R.; Perelygin, A.; Wu, J.; Chuang, J.; Manning, C.D.; Ng, A.; Potts, C. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2013, pp. 1631–1642.
5. Wu, S.; Fei, H.; Li, F.; Zhang, M.; Liu, Y.; Teng, C.; Ji, D. Mastering the Explicit Opinion-Role Interaction: Syntax-Aided Neural Transition System for Unified Opinion Role Labeling. Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence, 2022, pp. 11513–11521.
6. Shi, W.; Li, F.; Li, J.; Fei, H.; Ji, D. Effective Token Graph Modeling using a Novel Labeling Strategy for Structured Sentiment Analysis. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2022, pp. 4232–4241.
7. Fei, H.; Zhang, Y.; Ren, Y.; Ji, D. Latent Emotion Memory for Multi-Label Emotion Classification. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, pp. 7692–7699.
8. Wang, F.; Li, F.; Fei, H.; Li, J.; Wu, S.; Su, F.; Shi, W.; Ji, D.; Cai, B. Entity-centered Cross-document Relation Extraction. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, 2022, pp. 9871–9881.
9. Zhuang, L.; Fei, H.; Hu, P. Knowledge-enhanced event relation extraction via event ontology prompt. *Inf. Fusion* **2023**, *100*, 101919.
10. Udochukwu, O.; He, Y. A Rule-Based Approach to Implicit Emotion Detection in Text. NLDB, 2015.
11. Tan, L.I.; Phang, W.S.; Chin, K.O.; Patricia, A. Rule-Based Sentiment Analysis for Financial News. 2015 IEEE International Conference on Systems, Man, and Cybernetics, 2015, pp. 1601–1606. doi:10.1109/SMC.2015.283.
12. Fei, H.; Ren, Y.; Ji, D. Retrofitting Structure-aware Transformer Language Model for End Tasks. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020, pp. 2151–2161.
13. Seal, D.; Roy, U.; Basak, R., Sentence-Level Emotion Detection from Text Based on Semantic Rules; 2019; pp. 423–430. doi:10.1007/978-981-13-7166-0\_42.
14. Fei, H.; Wu, S.; Li, J.; Li, B.; Li, F.; Qin, L.; Zhang, M.; Zhang, M.; Chua, T.S. LasUIE: Unifying Information Extraction with Latent Adaptive Structure-aware Generative Language Model. Proceedings of the Advances in Neural Information Processing Systems, NeurIPS 2022, 2022, pp. 15460–15475.
15. Qiu, G.; Liu, B.; Bu, J.; Chen, C. Opinion word expansion and target extraction through double propagation. *Computational linguistics* **2011**, *37*, 9–27.
16. Fei, H.; Ren, Y.; Zhang, Y.; Ji, D.; Liang, X. Enriching contextualized language model from knowledge graph for biomedical information extraction. *Briefings in Bioinformatics* **2021**, *22*.
17. Wu, S.; Fei, H.; Ji, W.; Chua, T.S. Cross2StrA: Unpaired Cross-lingual Image Captioning with Cross-lingual Cross-modal Structure-pivoted Alignment. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 2593–2608.

18. Deng, Y.; Zhang, W.; Lam, W. Opinion-aware Answer Generation for Review-driven Question Answering in E-Commerce, 2020, [[arXiv:cs.CL/2008.11972](https://arxiv.org/abs/2008.11972)].
19. Chen, S.; Li, C.; Ji, F.; Zhou, W.; Chen, H. Review-Driven Answer Generation for Product-Related Questions in E-Commerce; Association for Computing Machinery: New York, NY, USA, 2019; WSDM '19, p. 411–419. doi:10.1145/3289600.3290971.
20. Wu, S.; Fei, H.; Qu, L.; Ji, W.; Chua, T.S. NEX-T-GPT: Any-to-Any Multimodal LLM. *CoRR* **2023**, *abs/2309.05519*.
21. Gao, S.; Ren, Z.; Zhao, Y.E.; Zhao, D.; Yin, D.; Yan, R. Product-Aware Answer Generation in E-Commerce Question-Answering. *CoRR* **2019**, *abs/1901.07696*, [[1901.07696](https://arxiv.org/abs/1901.07696)].
22. Zheng, S.; Xia, R. Left-Center-Right Separated Neural Network for Aspect-based Sentiment Analysis with Rotatory Attention, 2018, [[arXiv:1802.00892](https://arxiv.org/abs/1802.00892)]. *arXiv preprint arXiv:1802.00892*.
23. Truşcă, M.M.; Wassenberg, D.; Frasinca, F.; Dekker, R. A Hybrid Approach for Aspect-Based Sentiment Analysis Using Deep Contextual Word Embeddings and Hierarchical Attention. 20th International Conference on Web Engineering (ICWE 2020). Springer, 2020, Vol. 12128, LNCS, pp. 365–380.
24. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In *28th Annual Conference on Neural Information Processing Systems (NIPS 2014)*; Curran Associates, Inc., 2014; pp. 2672–2680.
25. Han, J.; Zhang, Z.; Schuller, B. Adversarial training in affective computing and sentiment analysis: Recent advances and perspectives. *IEEE Computational Intelligence Magazine* **2019**, *14*, 68–81.
26. Wu, S.; Fei, H.; Ren, Y.; Ji, D.; Li, J. Learn from Syntax: Improving Pair-wise Aspect and Opinion Terms Extraction with Rich Syntactic Knowledge. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, 2021, pp. 3957–3963.
27. Li, B.; Fei, H.; Liao, L.; Zhao, Y.; Teng, C.; Chua, T.; Ji, D.; Li, F. Revisiting Disentanglement and Fusion on Modality and Context in Conversational Multimodal Emotion Recognition. Proceedings of the 31st ACM International Conference on Multimedia, MM, 2023, pp. 5923–5934.
28. Fei, H.; Liu, Q.; Zhang, M.; Zhang, M.; Chua, T.S. Scene Graph as Pivoting: Inference-time Image-free Unsupervised Multimodal Machine Translation with Visual Scene Hallucination. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 5980–5994.
29. Karimi, A.; Rossi, L.; Prati, A.; Full, K. Adversarial training for aspect-based sentiment analysis with BERT. *arXiv preprint arXiv:2001.11316* **2020**.
30. Xu, H.; Liu, B.; Shu, L.; Yu, P.S. BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis. 2019 Conference of the North American Chapter of ACL: Human Language Technologies (NAACL-HLT 2019). ACL, 2019, pp. 2324–2335.
31. Li, J.; Xu, K.; Li, F.; Fei, H.; Ren, Y.; Ji, D. MRN: A Locally and Globally Mention-Based Reasoning Network for Document-Level Relation Extraction. Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, 2021, pp. 1359–1370.
32. Fei, H.; Wu, S.; Ren, Y.; Zhang, M. Matching Structure for Dual Learning. Proceedings of the International Conference on Machine Learning, ICML, 2022, pp. 6373–6391.
33. Cao, H.; Li, J.; Su, F.; Li, F.; Fei, H.; Wu, S.; Li, B.; Zhao, L.; Ji, D. OneEE: A One-Stage Framework for Fast Overlapping and Nested Event Extraction. Proceedings of the 29th International Conference on Computational Linguistics, 2022, pp. 1953–1964.
34. Wu, S.; Fei, H.; Cao, Y.; Bing, L.; Chua, T.S. Information Screening whilst Exploiting! Multimodal Relation Extraction with Feature Denoising and Multimodal Topic Modeling. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 14734–14751.
35. Fei, H.; Li, F.; Li, B.; Ji, D. Encoder-Decoder Based Unified Semantic Role Labeling with Label-Aware Syntax. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, pp. 12794–12802.
36. Li, B.; Fei, H.; Li, F.; Wu, Y.; Zhang, J.; Wu, S.; Li, J.; Liu, Y.; Liao, L.; Chua, T.S.; Ji, D. DiaASQ: A Benchmark of Conversational Aspect-based Sentiment Quadruple Analysis. Findings of the Association for Computational Linguistics: ACL 2023, 2023, pp. 13449–13467.
37. Pontiki, M.; Galanis, D.; Papageorgiou, H.; Manandhar, S.; Androutsopoulos, I. SemEval-2015 Task 12: Aspect Based Sentiment Analysis. 9th International Workshop on Semantic Evaluation (SemEval 2015). ACL, 2015, pp. 486–495.

38. Pontiki, M.; Galanis, D.; Papageorgiou, H.; Manandhar, S.; Androutsopoulos, I. SemEval-2016 Task 5: Aspect Based Sentiment Analysis. 10th International Workshop on Semantic Evaluation (SemEval 2016). ACL, 2016, pp. 19–30.
39. Fei, H.; Li, F.; Li, C.; Wu, S.; Li, J.; Ji, D. Inheriting the Wisdom of Predecessors: A Multiplex Cascade Framework for Unified Aspect-based Sentiment Analysis. Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI, 2022, pp. 4096–4103.
40. Odena, A. Semi-Supervised Learning with Generative Adversarial Networks. *arXiv preprint arXiv:1606.01583* 2016, [1606.01583].
41. Fei, H.; Wu, S.; Ren, Y.; Li, F.; Ji, D. Better Combine Them Together! Integrating Syntactic Constituency and Dependency Representations for Semantic Role Labeling. Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, 2021, pp. 549–559.
42. Wu, S.; Fei, H.; Zhang, H.; Chua, T.S. Imagine That! Abstract-to-Intricate Text-to-Image Synthesis with Scene Graph Hallucination Diffusion. *Advances in Neural Information Processing Systems* 2024, 36.
43. Fei, H.; Wu, S.; Ji, W.; Zhang, H.; Chua, T.S. Empowering dynamics-aware text-to-video diffusion with large language models. *arXiv preprint arXiv:2308.13812* 2023.
44. Qu, L.; Wu, S.; Fei, H.; Nie, L.; Chua, T.S. Layoutllm-t2i: Eliciting layout guidance from llm for text-to-image generation. Proceedings of the 31st ACM International Conference on Multimedia, 2023, pp. 643–654.
45. Fei, H.; Ren, Y.; Ji, D. Boundaries and edges rethinking: An end-to-end neural model for overlapping entity relation extraction. *Information Processing & Management* 2020, 57, 102311.
46. Li, J.; Fei, H.; Liu, J.; Wu, S.; Zhang, M.; Teng, C.; Ji, D.; Li, F. Unified Named Entity Recognition as Word-Word Relation Classification. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, pp. 10965–10973.
47. Abdul-Mageed, M.; Ungar, L. EmoNet: Fine-Grained Emotion Detection with Gated Recurrent Neural Networks. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); Association for Computational Linguistics: Vancouver, Canada, 2017; pp. 718–728. doi:10.18653/v1/P17-1067.
48. Fei, H.; Chua, T.; Li, C.; Ji, D.; Zhang, M.; Ren, Y. On the Robustness of Aspect-based Sentiment Analysis: Rethinking Model, Data, and Training. *ACM Transactions on Information Systems* 2023, 41, 50:1–50:32.
49. Zhao, Y.; Fei, H.; Cao, Y.; Li, B.; Zhang, M.; Wei, J.; Zhang, M.; Chua, T. Constructing Holistic Spatio-Temporal Scene Graph for Video Semantic Role Labeling. Proceedings of the 31st ACM International Conference on Multimedia, MM, 2023, pp. 5281–5291.
50. Fei, H.; Ren, Y.; Zhang, Y.; Ji, D. Nonautoregressive Encoder-Decoder Neural Framework for End-to-End Aspect-Based Sentiment Triplet Extraction. *IEEE Transactions on Neural Networks and Learning Systems* 2023, 34, 5544–5556.
51. Zhao, Y.; Fei, H.; Ji, W.; Wei, J.; Zhang, M.; Zhang, M.; Chua, T.S. Generating Visual Spatial Description via Holistic 3D Scene Understanding. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 7960–7977.
52. Raffel, C.; Shazeer, N.; Roberts, A.; Lee, K.; Narang, S.; Matena, M.; Zhou, Y.; Li, W.; Liu, P.J. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *CoRR* 2019, abs/1910.10683, [1910.10683].
53. Fei, H.; Li, B.; Liu, Q.; Bing, L.; Li, F.; Chua, T.S. Reasoning Implicit Sentiment with Chain-of-Thought Prompting. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2023, pp. 1171–1182.
54. Schouten, K.; Frasincar, F. Ontology-Driven Sentiment Analysis of Product and Service Aspects. 15th Extended Semantic Web Conference (ESWC 2018). Springer, 2018, Vol. 10843, LNCS, pp. 608–623.
55. Salimans, T.; Goodfellow, I.J.; Zaremba, W.; Cheung, V.; Radford, A.; Chen, X. Improved Techniques for Training GANs. 29th Annual Conference on Neural Information Processing Systems (NIPS 2016). Curran Associates, Inc., 2016, pp. 2226–2234.
56. Goodfellow, I.J. On distinguishability criteria for estimating generative models. 3rd International Conference on Learning Representations (ICLR 2015), Workshop Track, 2015.
57. Bird, S.; Klein, E.; Loper, E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*; O'Reilly Media, Sebastopol, CA, 2009.
58. Miller, G.A. WordNet: a lexical database for English. *Communications of the ACM* 1995, 38, 39–41.

59. Bergstra, J.; Bardenet, R.; Bengio, Y.; Kégl, B. Algorithms for Hyper-Parameter Optimization. 25th Annual Conference on Neural Information Processing Systems (NIPS 2011), 2011, pp. 2546–2554.
60. Li, Z.; Feng, C.; Zheng, J.; Wu, M.; Yu, H. Towards Adversarial Robustness via Feature Matching. *IEEE Access* **2020**, *8*, 88594–88603.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.