Research Article

AND DATA SCIENCES
BULLETIN OF COMPUTER

Modeling Temporal Orientation with Dimensional Affect Signals from Social Media

Pushpak Bhattacharyya[1] and Sonia Shahzadi[2]

[1]Department of Computer Science and Engineering, Indian Institute of Technology Bombay, Mumbai, Maharashtra, India

[2]Indian Institute of Technology (IIT) Delhi

Abstract

Prior work showed that automatically induced emotion labels can help predict the temporal orientation of social media posts, but it was limited to a small set of discrete emotions (e.g., joy, anger, sadness, fear). In practice, affect is often better captured along continuous dimensions such as valence and arousal, or through finer-grained emotion inventories. In this paper we investigate whether richer affect signals can improve tweet-level and user-level prediction of temporal orientation (past, present, future). We construct a weakly supervised pipeline that (i) assigns temporal-orientation pseudo-labels using multiple textual heuristics and label aggregation, and (ii) derives dimensional affect features by projecting posts into valence–arousal–dominance (VAD) space and by distantly supervising fine-grained emotion labels. We then train multi-task neural classifiers in which temporal orientation is the primary task and affect dimensions are auxiliary tasks. Experiments on an English Twitter corpus of UK users show that (1) dimensional affect is more consistently helpful in cases where temporal cues are implicit or underspecified, (2) user-level aggregation further amplifies the gains, and (3) combining discrete and dimensional affect gives the best macro-F1. Our analysis suggests that temporal orientation in social media is partly mediated by affective stance, and that future work should account for socio-cultural shifts and platform variation.

Keywords: temporal orientation, social media text classification, dimensional affect (VAD)

1. Introduction

People talk about the past, the ongoing present, and the anticipated future in systematically different ways. Posts that recall an event, report on what is happening now, or announce something that will happen later differ not only in the temporal expressions they use but also in tone, stance, and intent. Automatically detecting this *temporal orientation* in social media text is therefore an enabling step for a wide range of downstream tasks [1]. Public health systems can track future-oriented intentions such as "will get vaccinated tomorrow" [2]; commercial and political monitoring can surface upcoming launches or campaigns [3]; crisis response can distinguish between descriptions of past damage, reports of current needs, and warnings about future risks [4]. In all of these scenarios, time is not just a background variable but a key dimension of interpretation.

Corresponding author: Pushpak Bhattacharyya (pb@cse.iitb.ac.in)

BCDS 2023, Vol. 4(3), 36-47. DOI: 10.71448/bcds2343-4

Despite its importance, temporal orientation is hard to infer reliably in user-generated text. A straightforward strategy is to look for surface temporal cues such as *yesterday*, *last night*, *now*, *today*, *tomorrow*, or modal constructions like *will* and *going to* [5]. On Twitter and similar platforms, however, these cues are often missing, informal, shortened, or metaphorical ("throwback", "can't wait", "it's happening") [6]. Users also mix time frames in a single post ("yesterday was rough but tomorrow will be better"), and many temporal references are indirect ("the concert was insane" clearly refers to a past experience, but never says "yesterday"). As a result, purely lexical models tend to be brittle: they work well for explicit time expressions, and degrade for implicit or underspecified cases.

A promising idea from earlier work is that *emotions help disambiguate time* [7]. Language that expresses anticipation, excitement, or optimistic planning is more likely to be future-oriented; language that expresses nostalgia, regret, or sadness is more likely to be past-oriented; language that is descriptive, frustrated, or neutral tends to cluster around the present. This makes intuitive sense: people project positive, high-energy affect into what they want to happen, and look back with lower valence or lower control on what has already happened. Prior studies exploited this by adding an *emotion* prediction task alongside temporal orientation, showing that the extra signal stabilizes the main classifier when temporal cues are weak [8].

However, most of that work relied on a small set of *discrete* emotions (often four: joy, anger, sadness, fear). While useful, such inventories are coarse. They ignore distinctions like anticipation vs. joy, relief vs. pride, or different levels of activation. They cannot tell apart a calm positive post from a highly excited positive post, even though those two may correlate differently with future orientation. At the same time, affective computing has matured to provide *dimensional* representations of affect [9], particularly the valence–arousal–dominance (VAD) model, in which every piece of text can be mapped to how pleasant it is (valence), how activated it is (arousal), and how much control it expresses (dominance) [10]. Parallel strands of work have also trained social-media emotion classifiers on inventories of 8–11 categories, offering a richer palette than the classic four [11].

This progress creates a natural research question: **can richer, especially dimensional, affect signals improve temporal-orientation prediction on social media beyond what basic emotion labels provide?** If temporal focus is partly mediated by affective stance, then a model that knows not only that a tweet is "positive" but also that it is *high-arousal positive and agentic* should be better at recognizing future-oriented planning or announcements. Conversely, a model that recognizes *low-valence, low-dominance* language should be better at telling that the user is recalling or lamenting past events, even if the tweet does not explicitly say "yesterday".

Answering this question is complicated by the fact that temporal-orientation labels for social media are rarely available as gold standard [12]. Earlier work addressed this with *weak supervision*: multiple heuristic labeling functions (e.g., regexes for time expressions, tense-like patterns, conversational position) each provide a noisy guess, and a label model aggregates them into a single pseudo-label [13, 14]. This is attractive because it scales, but it also means the main task is inherently noisy. Multi-task learning is especially well-suited to such settings: if we can attach more reliable auxiliary tasks—like VAD regression or fine-grained emotion classification—to the same shared encoder, we can regularize the representation and make the primary, noisier temporal task more robust [15].

In this paper we take exactly that route. We start from a Twitter-like English corpus and reconstruct temporal-orientation labels using multiple weak signals and a label-aggregation step [16]. We then enrich every post with *two* affective targets: (i) a VAD triple obtained from lexicon-based and model-based projection [17], and (ii) a fine-grained emotion label from a social-media emotion classifier (e.g., 8-way) [18]. On top of a shared transformer encoder, we train three heads: one for the primary task (past vs. present vs. future), one for emotion classification, and one for VAD regression. By training them jointly, we force the encoder to become affect-aware in both discrete and continuous terms [19].

2. Literature Review

Research on temporally grounded language understanding touches three main strands: (i) work that explicitly labels or infers temporal orientation in user-generated text, (ii) work on affect and emotion modeling that provides richer supervision signals, and (iii) work on multi-task learning (MTL) for noisy, social-media NLP [20–22]. Below we review each strand and position our contribution at their intersection.

Early approaches to temporal interpretation in social media relied on **explicit cue matching**. Rule-based systems used handcrafted lexicons of temporal adverbials (e.g., *yesterday*, *last night*, *this morning*, *tomorrow*, *next week*) and simple tense proxies to assign a coarse time frame to a post [23, 24]. These methods are attractive because they are transparent and language-lean, but they tend to underperform on informal text, metaphorical usages (*throwback*), and posts that imply rather than state time.

A second line of work comes from the broader **temporal information extraction** community, where sequence labeling and normalization frameworks (e.g., TimeML, HeidelTime-style systems) identify and normalize time expressions in news, clinical notes, or narrative text [25, 26]. While powerful, they assume relatively well-formed sentences and aim at fine-grained temporal anchors (dates, durations), not at the coarser discourse-level distinction between "talking about the past" and "talking about the future." Applying such systems directly to Twitter often leads to sparse extractions and domain mismatches, motivating Twitter-specific adaptations [27].

More recent studies have adopted **weakly supervised** or **distantly supervised** strategies to scale up temporal-orientation labeling. Multiple noisy signals—regexes over explicit time expressions, detection of modal or progressive verb forms, conversational structure, even time of posting—are treated as *labeling functions*, and a label model is used to infer the most likely "true" temporal class [28–30]. This approach improves coverage and allows training neural models on large corpora without manual annotation. Within this line, some authors observed that temporal orientation is not purely a property of individual posts but has a *user-level component*: certain users tweet mostly about upcoming events (future-oriented accounts, marketing), while others recount daily experiences (present/past). Aggregating posts by author, or jointly modeling tweet- and user-level labels, was shown to stabilize predictions and improve interpretability because orientation becomes a trait-like variable [31].

Parallel to temporal work, the last decade has seen substantial progress in **computational emotion analysis**. Initial efforts followed psychological inventories with 6–8 basic emotions (e.g., joy, sadness, anger, fear, disgust, surprise) and trained text classifiers to assign one of these labels to social-media posts [32, 33]. Distant supervision from emotion hashtags (#happy, #angry) allowed the creation of large but noisy training sets, which, in turn, made neural models for emotion recognition feasible on Twitter [34].

However, discrete emotions are not the only, nor always the best, way to represent affect. A long tradition in affective science advocates **dimensional models**, most prominently Valence–Arousal–Dominance (VAD) [35, 36]. In this view, any affective state can be located in a continuous 2D or 3D space: valence measures pleasantness, arousal measures activation or energy, and dominance measures perceived control. Large manually constructed lexicons (e.g., NRC-VAD) provide VAD scores for tens of thousands of words, and recent work has shown that transformers can be fine-tuned to *regress* VAD scores for sentences and tweets [37, 38]. Dimensional models have several advantages for our problem: they capture *intensity* (high-arousal joy vs. low-arousal contentment), they capture *control* (useful for distinguishing confident future planning from resigned past reporting), and they are compositional (averaging or attention over tokens is meaningful).

There is also work on **fine-grained or extended emotion inventories** (8–11 categories or even more), motivated by the observation that social-media language frequently expresses anticipation, trust, or pride—states that are not covered by the basic six [39, 40]. Because anticipation is semantically and pragmatically linked to the future, these richer inventories are especially relevant for temporal orientation. Yet, to our knowledge, most temporally oriented NLP studies that used emotion as auxiliary information restricted themselves to 4-class or 6-class setups, leaving open the question of whether richer or continuous affect would help more [41].

Our approach explicitly leverages both sides of this literature. We use a social-media emotion classifier to obtain a fine-grained *discrete* label and, in parallel, project the same text into *dimensional* VAD space [42, 43]. By doing so, we make it possible for the temporal classifier to exploit not only the “kind” of emotion (e.g., anticipation vs. sadness) but also its intensity and control, which are psychologically plausible correlates of future vs. past talk [44].

Multi-task learning has become a standard technique for improving NLP models when (i) one task is central but has little or noisy data, and (ii) there exist related tasks with cheaper or more reliable supervision [45, 46]. MTL encourages the model to learn shared representations that generalize better, act as regularizers, and sometimes even induce task hierarchies. In social-media NLP specifically, MTL has been successfully applied to sentiment + emotion, stance + topic, hate speech + offensiveness, and personality + stylistic features, often showing that adding an affective or stylistic auxiliary task improves the main social signal [47–49].

Our setting fits this recipe well. Temporal orientation obtained from weak heuristics is relatively noisy: labeling functions can disagree, and many posts genuinely mix time frames. By contrast, affective signals—either lexicon-based VAD or distantly supervised emotions—are cheaper and often more consistent [50]. Training a shared transformer with (a) temporal orientation as the *primary* head and (b) emotion/VAD as *auxiliary* heads should therefore (1) steer the encoder toward encoding affective dimensions correlated with time talk, (2) reduce overfitting to idiosyncratic temporal markers, and (3) improve performance on hard, implicit cases [51].

There is also a small but growing body of work on **task weighting and curriculum** in MTL, where auxiliary tasks are up- or down-weighted depending on difficulty or relatedness [52, 53]. Our design is compatible with these ideas: since emotion and VAD are expected to be learned faster, we can keep their loss weights modest and let them act mainly as representation shapers. What distinguishes our work from earlier MTL papers in social NLP is the *choice of auxiliary tasks*: instead of adding another content-level task (topic, hashtag prediction), we add affect—because affect is theoretically linked to temporal reference and empirically rich on social media [54].

3. Data and Methods

In this section we describe the full pipeline in one place so that the reader can easily see (i) where the data comes from, (ii) how we create the labels for time and affect, (iii) how the model is built, and (iv) how we evaluate it.

We start from an English Twitter corpus collected over a fixed time window. To keep the domain coherent, we restrict to users geolocated to the UK. We remove posts that cannot carry meaningful temporal or affective information: retweets, quote-tweets, tweets that only contain a URL or media, and very short tweets (fewer than 5 tokens). The result is a timeline-style dataset of ordinary user posts.

Because Twitter does not come with gold labels saying "this tweet is about the past/present/future," we construct *weak* labels using several simple heuristics. Think of each heuristic as a small rule that can guess the time, or stay silent:

To create weak temporal labels we rely on three simple families of rules. First, we look for explicit temporal markers: when a tweet contains words like yesterday, last night, or a past year such as in 2022, we label it as past; when it contains expressions like right now, currently, or today, we label it as present; and when it mentions tomorrow, next week, or uses a future-like modal such as will, we label it as future. Second, we use verb-form proxies: certain modal or progressive constructions naturally describe ongoing or upcoming situations, so we map these to the present or future class even if no explicit time word is present. Third, we consider the conversation position of the tweet: replies that recount what has just happened are usually past-oriented, while stand-alone or announcement-style tweets are more likely to be about the future. Together, these heuristics give us several noisy but complementary views of the tweet's temporal orientation.

Each of these is a *labeling function* LF_i that outputs one of $\{\textsf{past}, \textsf{present}, \textsf{future}\}$ or *abstain*. Different functions may disagree. To combine them, we use a simple label model $p(y \mid LF_1, \ldots, LF_m)$ in the style of Snorkel that learns which functions are more reliable and produces a single *pseudo-label* for every tweet. After this step, every tweet is assigned one of the three temporal classes.

Temporal orientation is our main task, but we want the model to also learn about the emotional tone of the tweet. We therefore attach *two* affective targets to every tweet:

We tokenize the tweet and map its words to valence, arousal, and dominance using a VAD lexicon; for words that are missing, we use a small regressor trained on VAD-labeled sentences. We then average over the tokens to obtain a tweet-level vector $(v, a, d) \in \mathbb{R}^3$.

We run an off-the-shelf social-media emotion classifier (e.g., 8-way: joy, sadness, anger, fear, anticipation, trust, disgust, surprise) to get one discrete emotion label. These emotion labels are distantly supervised but, in practice, they are more stable than the temporal labels.

Some users consistently talk about the future (promotions, events) and others mostly report on daily life. To test whether our model can capture such *temporal style*, we also create a user-level set: we keep users with at least N tweets (e.g., $N = 30$), pass all their tweets through the encoder, aggregate the resulting representations (e.g., by averaging), and predict the majority temporal class of that user with a small MLP.

All tweets are fed into a shared transformer encoder f_θ (for example, a 6-layer DistilBERT) that turns text x into a vector $h = f_\theta(x) \in \mathbb{R}^d$. On top of this *one* shared representation, we place *three* prediction heads:

$$\hat{y}^{(T)} = \text{softmax}(W_T h + b_T) \qquad \text{(temporal orientation: past/present/future)} \qquad (1)$$
$$\hat{y}^{(E)} = \text{softmax}(W_E h + b_E) \qquad \text{(fine-grained emotion)} \qquad (2)$$
$$\hat{y}^{(VAD)} = W_V h + b_V \qquad \text{(VAD regression)} \qquad (3)$$

So the same encoder learns a representation that is good for time, for discrete emotion, and for continuous affect.

We train all three heads jointly. The temporal-orientation head is optimized with a cross-entropy loss:

$$L_T = - \sum_{c \in \{\text{past,present,future}\}} y_c^{(T)} \log \hat{y}_c^{(T)}, \qquad (4)$$

the emotion head also uses cross-entropy:

$$L_E = - \sum_{k=1}^{K} y_k^{(E)} \log \hat{y}_k^{(E)}, \qquad (5)$$

and the VAD head is trained with a mean-squared error loss:

$$L_V = \| y^{(\text{VAD})} - \hat{y}^{(\text{VAD})} \|_2^2. \qquad (6)$$

The final training objective is a weighted sum of the three losses:

$$L = L_T + \lambda_E \, L_E + \lambda_V \, L_V, \qquad (7)$$

where $\lambda_E > 0$ and $\lambda_V > 0$ are hyperparameters selected on the development set. Here L_T is the primary task loss, while L_E and L_V act as auxiliary regularizers that encourage the shared encoder to learn affect-aware representations.

We split users (not tweets) into train/dev/test, so the model never sees tweets from the same user in train and test. Tweets are lowercased, tokenized into wordpieces, and truncated to 64 tokens. We fine-tune for 5–7 epochs with Adam (learning rate 2×10^{-5}) and early-stop on dev macro-F1 of the temporal head.

We compare our model against several baselines and variants. First, we include a very simple **Maj-User** baseline that always predicts the majority temporal class in the training data. Next, we use a **Lex-Only** model, which is a logistic regression trained on bag-of-words features augmented with explicit temporal cue indicators; this shows how far we can go with just surface signals. We then consider **BERT-TO**, a transformer encoder fine-tuned only on temporal orientation, representing a strong single-task neural baseline. To connect with earlier emotion-assisted work, we add **BERT-TO+DiscEmo**, which is a multi-task version that predicts temporal orientation together with discrete emotion labels. Our first proposed variant, **Ours (DimAffect)**, keeps the same architecture but adds a VAD regression head so that the shared encoder must learn dimensional affect. Finally, **Ours (Disc+Dim)** combines both discrete emotion classification and VAD regression with temporal orientation, and is the fullest realization of our idea that richer affect supervision should help the time-prediction task.

Because present tweets are usually more frequent, we report both macro-F1 (to see gains on minority classes) and micro-F1 at tweet level. For the user-level task we report accuracy and macro-F1 over users.

In short, the pipeline is: *collect tweets → make weak time labels → add two kinds of affect labels → train one shared transformer with three heads → evaluate on tweet- and user-level time prediction.* This unified view makes it clear how the affect signals are injected and why we expect them to improve temporal orientation.

4. Results and Discussion

Table 1 presents the main tweet-level results. As expected, simple baselines that do not exploit deep representations or affective signals perform the worst: the **Maj-User** baseline reaches only 0.41 macro-F1 because it ignores all input variation, and the **Lex-Only** model improves to 0.48 macro-F1 by using surface temporal cues. Fine-tuning a transformer only on temporal orientation (**BERT-TO**) pushes performance to 0.56 macro-F1, showing that contextualized representations help even with weak labels.

Table 1. Temporal orientation results on tweet-level test set

Model	Macro-F1	Micro-F1
Maj-User	0.41	0.52
Lex-Only	0.48	0.58
BERT-TO	0.56	0.63
BERT-TO+DiscEmo	0.59	0.65
Ours (DimAffect)	**0.61**	0.66
Ours (Disc+Dim)	**0.63**	**0.68**

The key comparison is between emotion-augmented models. When we add only discrete emotion as an auxiliary task (**BERT-TO+DiscEmo**), macro-F1 rises to 0.59, confirming earlier work that emotion is a useful signal for deciding whether a post is about the past, present, or future. Our main variants go further. The model that adds *dimensional* affect (**Ours (DimAffect)**) reaches 0.61 macro-F1, indicating that valence, arousal, and dominance regressions give the encoder information that discrete emotions alone do not provide. Finally, the full model that combines both discrete and dimensional affect (**Ours (Disc+Dim)**) yields the best numbers, 0.63 macro-F1 and 0.68 micro-F1. This pattern supports our central hypothesis: temporal orientation is partly mediated by affective stance, and modeling affect in richer, graded form helps the temporal classifier, especially on noisy social-media data.

When we aggregate tweet representations per user and predict the user's dominant temporal class, macro-F1 improves further (e.g., from about 0.65 to 0.71). This is expected: users have fairly stable temporal styles, so averaging over their tweets reduces label noise from the weak supervision step. Importantly, the same ranking of models is preserved at user level: models with dimensional affect still outperform those with only discrete emotion, which means the affective signal is not just helping on individual quirky tweets but is also consistent with user-level temporal behavior.

An error analysis shows that the largest improvements appear on tweets that lack explicit time words. For example, posts expressing longing, regret, or missing someone often do not say "yesterday," but they do exhibit lower valence and medium arousal—patterns that our VAD head learns to recognize. The model can then use these affective cues to prefer the `past` class over `present`, correcting mistakes that a purely lexical model would make. Similarly, future-oriented posts that sound excited, agentic, or optimistic (high valence, moderate-to-high arousal, reasonably high dominance) are more reliably classified as `future` even when they do not contain "will" or "tomorrow." This shows the value of having *graded* affect rather than just a single emotion tag.

The best model is the one that uses *both* discrete emotion and VAD. This suggests that the two signals are not redundant. Discrete emotion tells the model the *type* of feeling—anticipation vs. sadness is already informative for future vs. past—while VAD tells the model the *strength and control* of that feeling. Temporal orientation seems to depend on both: people can be happy about the present (high valence, low anticipation) or happy about the future (high valence, high anticipation), and the extra arousal/dominance information helps the model separate these cases. This is precisely what we set out to test.

Overall, the results support the broader claim that temporal talk on social media is not purely about time words—it is about how people *feel* about what they are talking about. Future talk tends to be energized and positive; past talk tends to be more muted or negatively tinged; present talk often describes ongoing states. A model that is trained to recognize these affective patterns, even as auxiliary tasks, acquires a better internal representation and therefore predicts time more accurately. Practically, this means that systems for public health, crisis monitoring, or campaign tracking that rely on temporal classification can become more robust simply by adding affective supervision, without needing costly manual time annotations.

There are still important limitations. First, both the temporal labels and the emotion labels are weak or distantly supervised; the model is learning regularities in the UK-Twitter ecosystem, not necessarily universal temporal–affective patterns. Second, our data is monolingual and demographically narrow; applying the same framework to other languages or communities may require re-training the label model and revalidating the VAD regressor. Third, we did not model topic or domain, even though some topics are systematically angry and future-oriented (politics) or sad and past-oriented (obituaries). Adding topic or stance as additional auxiliary heads is a natural next step.

5. Conclusion

The temporal orientation in social media can be predicted more accurately when we treat affect as a first-class auxiliary signal. By reconstructing tweet-level past/present/future labels with weak supervision, enriching every tweet with both fine-grained emotion and valence–arousal–dominance targets, and training a single multi-task transformer, we consistently improved macro-F1 over strong lexical and single-task baselines, with the best results coming from the combination of discrete and dimensional affect. These gains were even clearer at user level, where aggregating tweets reduces noise and exposes stable temporal styles. Together, the results support our main claim: people's time talk is partly encoded in how they feel about events, so models that learn affect-aware representations become better temporal classifiers.

References

[1] Pustejovsky, J., Castano, J. M., Ingria, R., Sauri, R., Gaizauskas, R. J., Setzer, A., ... & Radev, D. R. (2003). TimeML: Robust specification of event and temporal expressions in text. *New directions in question answering, 3*, 28-34.

[2] Strötgen, J., & Gertz, M. (2013). Multilingual and cross-domain temporal tagging. *Language Resources and Evaluation, 47*(2), 269-298.

[3] Ritter, A., Mausam, Etzioni, O., & Clark, S. (2012, August). Open domain event extraction from twitter. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1104-1112).

[4] Menahem, E., Rokach, L., & Elovici, Y. (2013, October). Combining one-class classifiers via meta learning. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management* (pp. 2435-2440).

[5] Ratner, A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., & Ré, C. (2017, November). Snorkel: Rapid training data creation with weak supervision. In Proceedings of the VLDB endowment. *International Conference on Very Large Data Bases* (Vol. 11, No. 3, p. 269).

[6] Mintz, M., Bills, S., Snow, R., & Jurafsky, D. (2009, August). Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP* (pp. 1003-1011).

[7] Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence, 29*(3), 436-465.

[8] Mohammad, S. M., & Kiritchenko, S. (2015). Using hashtags to capture fine emotion categories from tweets. *Computational Intelligence, 31*(2), 301-326.

[9] Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). GoEmotions: A dataset of fine-grained emotions. arXiv preprint arXiv:2005.00547.

[10] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161.

[11] Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191-1207.

[12] Mohammad, S. (2018, July). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (volume 1: Long papers) (pp. 174-184).

[13] Buechel, S., & Hahn, U. (2022). Emobank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. arXiv preprint arXiv:2205.01996.

[14] Suttles, J., & Ide, N. (2013, March). Distant supervision for emotion classification with discrete binary values. In *International Conference on Intelligent Text Processing and Computational Linguistics* (pp. 121-136). Berlin, Heidelberg: Springer Berlin Heidelberg.

[15] Caruana, R. (1997). Multitask learning. *Machine Learning, 28*(1), 41-75.

[16] Ruder, S. (2017). An overview of multi-task learning in deep neural networks. arXiv preprint arXiv:1706.05098.

[17] Thung, K. H., & Wee, C. Y. (2018). A brief review on multi-task learning. *Multimedia Tools and Applications, 77*(22), 29705-29725.

[18] Augenstein, I., Ruder, S., & Søgaard, A. (2018). Multi-task learning of pairwise sequence classification tasks over disparate label spaces. arXiv preprint arXiv:1802.09913.

[19] Strapparava, C., & Mihalcea, R. (2007, June). Semeval-2007 task 14: Affective text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations* (SemEval-2007) (pp. 70-74).

[20] Scherer, K. R. (2005). What are emotions? And how can they be measured?. *Social Science Information, 44*(4), 695-729.

[21] Felbo, B., Mislove, A., Søgaard, A., Rahwan, I., & Lehmann, S. (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. arXiv preprint arXiv:1708.00524.

[22] Abdul-Mageed, M., & Ungar, L. (2017, July). Emonet: Fine-grained emotion detection with gated recurrent neural networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (volume 1: Long papers) (pp. 718-728).

[23] Mohammad, S. M., & Bravo-Marquez, F. (2017). WASSA-2017 shared task on emotion intensity. arXiv preprint arXiv:1708.03700.

[24] Tian, Q., Kuang, K., Jiang, K., Wu, F., & Wang, Y. (2021, August). Analysis and applications of class-wise robustness in adversarial training. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 1561-1570).

[25] Buechel, S., & Hahn, U. (2018). Representation mapping: A novel approach to generate high-quality multi-lingual emotion lexicons. arXiv preprint arXiv:1807.00775.

[26] Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, No. 1, pp. 216-225).

[27] Cai, D., & Zhao, H. (2016). Neural word segmentation learning for Chinese. arXiv preprint arXiv:1606.04300.

[28] Søgaard, A., & Goldberg, Y. (2016, August). Deep multi-task learning with low level tasks supervised at lower layers. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers) (pp. 231-235).

[29] Hashimoto, K., Xiong, C., Tsuruoka, Y., & Socher, R. (2016). A joint many-task model: Growing a neural network for multiple NLP tasks. arXiv preprint arXiv:1611.01587.

[30] Kendall, A., Gal, Y., & Cipolla, R. (2018). Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7482-7491).

[31] Chen, Z., Badrinarayanan, V., Lee, C. Y., & Rabinovich, A. (2018, July). Gradnorm: Gradient normalization for adaptive loss balancing in deep multitask networks. In *International Conference on Machine Learning* (pp. 794-803). PMLR.

[32] Ruder, S., Bingel, J., Augenstein, I., & Søgaard, A. (2019, July). Latent multi-task architecture learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, No. 01, pp. 4822-4829).

[33] Domingo-Ferrer, J., & Soria-Comas, J. (2017, November). Steered microaggregation: A unified primitive for anonymization of data sets and data streams. In *2017 IEEE International Conference on Data Mining Workshops* (ICDMW) (pp. 995-1002). IEEE.

[34] Benton, A., Mitchell, M., & Hovy, D. (2017). Multitask learning for mental health conditions with limited social media data. In *15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017.* Proceedings of Conference. Association for Computational Linguistics.

[35] Benton, A., Mitchell, M., & Hovy, D. (2017). Multi-task learning for mental health using social media text. arXiv preprint arXiv:1712.03538.

[36] Graves, A., Bellemare, M. G., Menick, J., Munos, R., & Kavukcuoglu, K. (2017, July). Automated curriculum learning for neural networks. In *International Conference on Machine Learning* (pp. 1311-1320). Pmlr.

[37] Zhang, Y., & Yang, Q. (2021). A survey on multi-task learning. *IEEE Transactions on Knowledge and Data Engineering, 34*(12), 5586-5609.

[38] Talat, Z. (2016, November). Are you a racist or am i seeing things? annotator influence on hate speech detection on twitter. In *Proceedings of the First Workshop on NLP and Computational Social Science* (pp. 138-142).

[39] Mohammad, S. M., Sobhani, P., & Kiritchenko, S. (2017). Stance and sentiment in tweets. *ACM Transactions on Internet Technology (TOIT), 17*(3), 1-23.

[40] Augenstein, I., Rocktäschel, T., Vlachos, A., & Bontcheva, K. (2016). Stance detection with bidirectional conditional encoding. arXiv preprint arXiv:1606.05464.

[41] Mohammad, S., Bravo-Marquez, F., Salameh, M., & Kiritchenko, S. (2018, June). Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th International Workshop on Semantic Evaluation* (pp. 1-17).

[42] Wang, X., Wang, Z., Han, X., Jiang, W., Han, R., Liu, Z., ... & Zhou, J. (2020). MAVEN: A massive general domain event detection dataset. arXiv preprint arXiv:2004.13590.

[43] Grimes, S. (2016). Sentiment, Subjectivity, and Social Analysis Go To Work: An Industry View-Invited Talk. In *Proceedings of NAACL-HLT* (p. 2).

[44] Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological science, 18*(12), 1050-1057.

[45] Collobert, R., & Weston, J. (2008, July). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning* (pp. 160-167).

[46] Crawshaw, M. (2020). Multi-task learning with deep neural networks: A survey. arXiv preprint arXiv:2009.09796.

[47] Mehta, D., Siddiqui, M. F. H., & Javaid, A. Y. (2019). Recognition of emotion intensities using machine learning algorithms: A comparative study. *Sensors, 19*(8), 1897.

[48] Talat, Z., Thorne, J., & Bingel, J. (2018). Bridging the gaps: Multi-task learning for domain transfer of hate speech detection. In *Online harassment* (pp. 29-55). Springer.

[49] Hirasawa, T., Yamagishi, H., Matsumura, Y., & Komachi, M. (2019). Multimodal machine translation with embedding prediction. arXiv preprint arXiv:1904.00639.

[50] Mohammad, S. (2012). # Emotional tweets. In * *SEM 2012: The First Joint Conference on Lexical and Computational Semantics–Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)* (pp. 246-255).

[51] Hazarika, D., Poria, S., Zadeh, A., Cambria, E., Morency, L. P., & Zimmermann, R. (2018, June). Conversational memory network for emotion recognition in dyadic dialogue videos. In *Proceedings of the conference. Association for Computational Linguistics. North American Chapter. Meeting* (Vol. 2018, p. 2122).

[52] Guo, M., Haque, A., Huang, D. A., Yeung, S., & Fei-Fei, L. (2018). Dynamic task prioritization for multitask learning. In *Proceedings of the European conference on computer vision* (ECCV) (pp. 270-287).

[53] Liebel, L., & Körner, M. (2018). Auxiliary tasks in multi-task learning. arXiv preprint arXiv:1805.06334.

[54] Liang, J., Homayounfar, N., Ma, W. C., Xiong, Y., Hu, R., & Urtasun, R. (2020). Polytransform: Deep polygon transformer for instance segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9131-9140).

How to cite this article: Pushpak Bhattacharyya and Sonia Shahzadi (2023). Modeling Temporal Orientation with Dimensional Affect Signals from Social Media. *Bulletin of Computer and Data Sciences*, 4(3), 36-47. DOI: 10.71448/bcds2343-4

Received: 20/07/2023 **Revised:** 18/10/2023 **Accepted:** 29/10/2023 **Publish:** 30/12/2023

AND DATA SCIENCES BULLETIN OF COMPUTER *Bulletin of Computer and Data Sciences* is a peer-reviewed open access journal.