

# Tinjauan Literatur : *Named Entity Recognition* pada Ulasan Wisata

Muhammad Fakhri Despawida Aulia Putra  
Program Studi Sarjana Informatika  
Universitas Islam Indonesia  
Jl. Kaliurang KM 14.5, Sleman, Yogyakarta, Indonesia  
17523177@student.uii.ac.id

Ahmad fathan Hidayatullah  
Program Studi Sarjana Informatika  
Universitas Islam Indonesia  
Jl. Kaliurang KM 14.5, Sleman, Yogyakarta, Indonesia  
Fathan@uui.ac.id

**Abstract**— Indonesia memiliki banyak sekali tempat wisata, membuat orang-orang antusias untuk mengunjunginya. Beberapa orang memilih untuk melihat ulasan wisata terlebih dahulu sebagai pertimbangan. Tetapi banyaknya tempat wisata membuat orang-orang kesulitan menemukan informasi yang diinginkan dari ulasan tersebut. *Named Entity Recognition* (NER) berguna untuk mengekstraksi informasi pada sebuah teks sehingga dapat lebih memudahkan orang-orang untuk mengetahui informasi yang terdapat pada suatu teks maupun dokumen. NER dapat digunakan untuk mempermudah menemukan informasi yang diinginkan seperti nama wisata, nama lokasi dan fasilitas. NER sudah diterapkan pada bidang biomedis, berita, medis, *twitter* dan *tourism*. Metode yang digunakan dalam NER adalah LSTM, BiLSTM, CNN dan CRF. Sedangkan penelitian NER pada domain *tourism* masih sedikit dilakukan dan masih harus dikembangkan lagi. Hasil penelitian ini diharapkan bisa digunakan untuk membantu mengembangkan penelitian NER selanjutnya. Tinjauan literatur ini dibuat untuk mengkaji literatur sebelumnya tentang NER dengan *deep learning* dan NER pada bidang *tourism* sehingga dapat membantu pengembangan penerapan NER selanjutnya

**Keywords**—*Named Entity Recognition*, *deep learning*, *ekstraksi fitur*

## I. PENDAHULUAN

Indonesia adalah negara yang salah satunya dipenuhi dengan banyak sekali tempat wisata. Dengan banyaknya tempat wisata yang terdapat di Indonesia ini membuat banyak orang antusias mengunjungi tempat tersebut. Akan tetapi dengan banyaknya tempat wisata, sering kali membuat orang-orang kebingungan untuk memilih destinasi wisata. Dikutip dari halaman *website* TripAdvisor, 83% pengguna global TripAdvisor biasanya selalu melakukan *review* pada ulasan-ulasan di TripAdvisor sebelum membuat keputusan. Pengguna juga setuju TripAdvisor membantu dalam merencanakan perjalanan wisata dengan baik dan mereka memilih untuk mempertimbangkan serta membaca ulasan untuk merencanakan perjalanan wisata [1]. Akan tetapi dengan banyaknya ulasan yang ada dapat membuat orang-orang kesulitan dalam menemukan informasi yang diinginkan.

*Named Entity Recognition* (NER) dapat berguna untuk mengekstraksi informasi pada sebuah teks dengan mengidentifikasi dan mengenali entitas yang ada [2]. *Named Entity Recognition* dapat digunakan untuk membantu pengguna dalam mengetahui informasi penting yang dibutuhkan. Dalam masalah ini, informasi yang dibutuhkan seperti nama wisata, nama lokasi dan fasilitas, sehingga dapat membantu pengguna dalam mendapatkan informasi tersebut.

Oleh karena itu, tinjauan pustaka ini ditulis bertujuan untuk melakukan komparasi metode pada NER, seperti metode apa saja yang sudah digunakan pada penelitian sebelumnya dan bagaimana hasilnya. Selain itu, bertujuan juga untuk melihat tren penerapan dari NER, untuk bidang apa saja NER diterapkan. Tinjauan literatur ini juga dilakukan untuk mengetahui sejauh mana penelitian NER dilakukan dalam bidang *tourism*. Penelitian ini diharapkan dapat memiliki kontribusi membantu peneliti dalam menentukan metode yang tepat untuk NER dan mengetahui tren penelitian dalam bidang NER.

## II. METODOLOGI PENELITIAN

### A. Pertanyaan Penelitian

Langkah pertama adalah mengidentifikasi pertanyaan penelitian. Pertanyaan yang dilakukan harus ringkas dan jelas. Pada penelitian ini, pertanyaan penelitiannya adalah sebagai berikut :

- (i) Bagaimana perkembangan tren penelitian yang digunakan dalam NER?
- (ii) Apa metode yang digunakan dalam NER?
- (iii) Bagaimana penelitian NER di bidang *tourism* pada penelitian sebelumnya?

### B. Pengumpulan Data

Pengumpulan data mencari literatur penelitian sebelumnya dilakukan menggunakan *Google Scholar*, *Medwell*, *Elsevier*, *Research Gate* dan *arXiv*. Dalam tahap ini literatur yang dikumpulkan merupakan penelitian NER dalam bidang umum yang menggunakan *deep learning* dan juga penelitian NER pada bidang *tourism*. Tinjauan literatur ini menganalisis literatur mulai dari tahun 2015 hingga 2020. Pengumpulan data dalam bidang umum ini menggunakan kata kunci pencarian "*Named Entity Recognition with Deep Learning*". Sedangkan untuk penelitian NER bidang *tourism* menggunakan kata kunci pencarian "*Named Entity Recognition Tourism*". Literatur yang berhasil dikumpulkan yaitu sebanyak 15 jurnal dengan 11 literatur membahas tentang umum dan 4 literatur membahas tentang *tourism*. Seperti pada tabel 1, berisi daftar mengenai referensi dan tahunnya.

TABLE I. TABEL REFERENSI DAN TAHUN PENELITIAN

No.	Tahun	Referensi
1.	2018	[3]

2.	2017	[4]
3.	2017	[5]
4.	2018	[2]
5.	2018	[6]
6.	2019	[7]
7.	2020	[8]
8.	2016	[9]
9.	2016	[10]
10.	2019	[11]
11.	2016	[12]
12.	2017	[13]
13.	2019	[14]
14.	2015	[15]
15.	2019	[16]

### III. PEMBAHASAN

Pada bab ini akan dibahas mengenai hasil kajian dan pembahasan serta temuan dari literatur tentang NER.

#### A. Hasil Kajian Literatur

Setelah dilakukan kajian dari literatur, diketahui bahwa dalam melakukan NER memiliki beberapa tahapan. Dalam bagian ini akan membahas bagaimana tahapan-tahapan dalam melakukan NER. Tahapan tersebut terdiri dari beberapa langkah seperti pengumpulan data NER, *preprocessing*, ekstraksi fitur dan juga penerapan model.

##### 1) Pengumpulan Data

Data yang digunakan dalam NER berupa teks atau dokumen yang didapatkan dari berbagai macam sumber. Seperti pada tabel 3 yang menampilkan *dataset* dari penelitian sebelumnya yang mencakup jumlah datanya dan juga jenis bahasanya yang diambil dalam penelitian. Jika dilihat dari tabel 2 tersebut maka dapat dianalisis bahwa dari 15 literatur penelitian, bahasa yang paling banyak digunakan adalah bahasa Inggris yang berjumlah 7 dan bahasa Indonesia yang berjumlah 5.

TABLE II. TABEL REFERENSI DATASET, JUMLAH DAN BAHASA PADA PENELITIAN

Referensi	Dataset	Jumlah	Bahasa
[3]	China EMR	55.485 kalimat	Cina
[4]	Twitter	480 tweet	Indonesia
[5]	GM BioCreative II dan korpus JNLPBA	20.000 dan 22.402 kalimat	Inggris
[2]	Wikipedia	700.000 artikel	Indonesia
[6]	Twitter dan Gazetteer	1.152 tweet	Indonesia

	OpenStreetMap (OSM)		
[7]	Artikel Web	-	Indonesia
[8]	China Judgements Online	1.000 dokumen	Cina
[9]	ConLL 2003 dan OntoNotes 5.0	5.000 kalimat	Inggris dan Jerman
[10]	TripAdvisor dan Wikipedia	6.996 kalimat	Inggris
[11]	TripAdvisor, Traveloka, Hotel.com	-	Inggris
[12]	Google	Ribuan	Inggris
[13]	i2b2	826 catatan	Inggris
[14]	Web Berita Indonesia	8.497 kalimat	Indonesia
[15]	PUBMED Database	339.084 makalah	Inggris
[16]	Ctrip, Mafengwo, Raiders	13.464 kalimat	Cina

##### 2) Preprocessing

*Preprocessing* merupakan tahapan untuk memproses dan mempersiapkan data agar lebih mudah untuk diolah. Dalam tahap ini ada beberapa jenis *preprocessing* yang dapat digunakan seperti *case folding*, normalisasi, tokenisasi dan *stopword*, bergantung sesuai dengan kebutuhannya. Pada penelitian [8], mereka menghapus seluruh spasi dan memberi keterangan pada data dengan label BIO. Sedangkan [6], melakukan beberapa *preprocessing* dengan menghapus semua tanda baca kecuali tanda tanya (?), titik (.) dan strip (-) kemudian menghapus kata "RT" di awal kalimat, menghapus URL, *case folding* dan mengganti kata singkatan dengan nama aslinya.

##### 3) Ekstraksi Fitur

Tahapan ekstraksi fitur digunakan untuk mengubah data mentah yang berupa teks menjadi vektor, ini dilakukan karena *deep learning* tidak dapat bekerja secara langsung pada data mentah. Melalui ekstraksi fitur ini, algoritma pembelajaran dapat mengetahui karakter dari sebuah data. *Word vector representation* atau yang biasa disebut *word embedding* dapat melakukan konversi sebuah teks menjadi angka. Bererapa macam *word embedding* yang dapat digunakan seperti *Word2Vec*, *GloVe* dan *Skip-gram*.

Pada penelitian [2], *word embedding* dilatih dengan model ruang vektor menggunakan pendekatan *Skip-gram*. Representasi kata yang terdistribusi pada ruang vektor berguna untuk membantu algoritma pembelajaran dalam mendapatkan performa yang lebih baik dengan mengelompokkan kata yang mirip. Kemudian [4] menggunakan *word embedding*, *neighbor word embedding* dan *POS Tag*. Hasil dari *word embedding* akan menjadi input

pada *neighbor word embedding*. *Neighbor word embedding* ini terdiri dari satu kata di sebelah kiri dan juga satu kata di sebelah kanan kata saat ini, jika kata ada di awal kalimat maka kata tidak memiliki kata di sebelah kiri sehingga vektor di sebelah kiri adalah 0, jika kata ada di akhir kalimat maka sebaliknya bahwa vektor 0 di sebelah kanan. Untuk *POS Tag*, ini dilakukan untuk melabeli tiap kata dengan kata kerja, kata benda, kata sifat dan sebagainya. Selain itu [4] menerapkan *propose the continuous bag-of-words* (CBOW) dari kumpulan data yang tidak dianotasi karena teks biomedis berbeda dengan korpora domain umum.

#### 4) Model NER

Dalam penelitian NER sebelumnya, terdapat beberapa model metode yang digunakan dalam membangun NER. Metode-metode yang telah digunakan sebelumnya dapat dilihat pada tabel 3.

TABLE III. TABEL METODE DAN REFERENSI

Referensi	Metode
[4]	LSTM, BiLSTM
[5]	LSTM, BiLSTM
[2]	BiLSTM, CNN
[6]	Recurrent CNN
[7]	BiLSTM, CRF
[8]	BiLSTM, CRF
[9]	BiLSTM, CNN
[3]	Multitask BiRNN
[11]	BERT
[10]	CRF
[12]	Naive Bayes Classifier
[13]	RNN, CNN, CRF
[14]	BiLSTM, CRF
[15]	CNN
[16]	BERT, BiLSTM, CRF

#### B. Pembahasan

Setelah dilakukan analisis dari seluruh literatur penelitian yang ada, ditemukan bahwa perkembangan tren penelitian dalam penerapan NER menggunakan *deep learning* lebih banyak menggunakan metode BiLSTM yang dikombinasikan dengan metode lain seperti LSTM, CNN dan CRF serta dapat ditambahkan dengan beberapa fitur seperti *word embeddings* dan *POS Tag*. BiLSTM banyak digunakan karena dapat memahami dan mengambil informasi terdahulu dan informasi terdahulu, sehingga proses pembelajaran akan semakin dalam yang berdampak pada model akan lebih memahami konteks pada ekstraksi informasi tersebut sehingga BiLSTM dapat melabeli entitas secara sekuensial apabila memiliki akses terhadap kedua informasi baik dari informasi sebelumnya dan sesudahnya. Kemudian dalam penelitian NER sebelumnya juga diketahui bahwa penelitian telah dilakukan di berbagai bidang. Bidang ataupun domain dalam NER yaitu pada media sosial *Twitter* yang menganalisis *tweet* bahasa Indonesia, seperti yang dilakukan oleh [4]. Selain itu [6] juga melakukan NER yang sama tetapi untuk mengenali sebuah kejadian dari sebuah *tweet*.

Pada bidang biomedis terdapat penelitian yang dilakukan [5], dengan mengenali entitas dalam domain

biomedis seperti *DNA*, *RNA*, protein, *cell line*, dan *cell type*. Selain itu [15] juga melakukan penelitian dalam domain biomedis dan mengenali entitas protein, *DNA*, *RNA*, *cell type* dan *cell line*. Dalam bidang klinis penelitian dilakukan oleh [13] dengan menerapkan NER untuk mengenali entitas pada catatan klinis seperti *problem*, *treatment* dan *test*. Penelitian yang dilakukan [2] menerapkan NER dalam domain berita dengan mengambil berita dan artikel tentang sejarah Indonesia. Domain yang sama juga dilakukan oleh [7] tetapi berita yang diambil adalah berita tentang politik. Penelitian [9] juga melakukan NER pada artikel berita, akan tetapi tidak memiliki kriteria khusus tentang data berita apa yang akan diambil. Penelitian oleh [14] juga melakukan penelitian dalam domain berita berbahasa Indonesia tetapi juga tidak memiliki kriteria khusus tentang berita yang diambil serta mengidentifikasi entitas seperti nama orang, nama lokasi, nama produk atau *brand*, nama *event* dan nama *beverages*. Kemudian penelitian [8] melakukan NER dalam bidang hukum dengan mengidentifikasi entitas *name*, *organization*, *judicial organization*, *docket number*, dan *crime type*. Dalam bidang medis, penelitian dilakukan oleh [3] dengan menerapkan NER pada *electronic medical records* yang berbahasa Cina.

Selain itu untuk bidang *tourism* sendiri, baru empat penelitian yang ditemukan dan tidak menggunakan *deep learning* karena minimnya penelitian. Penelitian NER bidang *tourism* telah dilakukan oleh [11], [10], [12] dan [16]. Penelitian [11] mengambil data dari artikel web yaitu TripAdvisor, Traveloka, dan *Hotels.com*. dengan bahasa Inggris. Penelitian ini menggunakan BERT dan spaCy dalam melakukan NER, sedangkan untuk entitas yang diidentifikasi adalah *location*, *organization* dan *facility*. Sedangkan [10] mengambil data dari TripAdvisor dan *Wikipedia* dengan mengidentifikasi entitas *location*, *person*, *organization*, *money*, *percent*, *date* dan *time* serta menggunakan metode CRF dalam melakukan NER. Kemudian [12] mengambil ribuan data *tourism* dari *google* dengan mengenali entitas *nature*, *place*, *city*, *region* dan *negative class*, metode *K-Nearest Neighbor* (KNN) dan *Naive Bayes Classifier* (NBC) dipilih untuk melakukan NER dalam penelitian ini. Setelah dilihat, diketahui bahwa dalam penerapan NER pada bidang *tourism* telah digunakan beberapa metode seperti CRF, BERT, KNN dan NBC. Belum ditemukan penelitian NER yang menggunakan *deep learning*. Selain itu pada literatur NER *tourism*, entitas yang dikenali berupa *person*, *location*, *organization*, *money*, *percent*, *date*, *time*, *facility*, *nature*, *place*, *city*, *region*, dan *negative class*. Penelitian oleh [16] melakukan NER dengan data yang diambil dari artikel wisata seperti *Ctrip*, *Mafengwo* dan *Raiders*. Penelitian ini menggunakan metode BBLC yaitu kombinasi BERT, BiLSTM dan CRF, mengenali entitas seperti *name*, *location*, *organization*, *time*, dan *thing*. Oleh karena itu dapat diketahui bahwa pada literatur NER *tourism* entitas yang paling sering diidentifikasi adalah nama, lokasi dan organisasi. Pada penelitian NER *tourism* data yang diambil juga masih terbatas dan hanya dari *website*.

Temuan selanjutnya yaitu tentang hasil dari penelitian yang sudah ada tentang NER. Dapat diketahui bahwa seperti yang ada pada tabel 4, diidentifikasi dengan referensi, hasil penelitian dan entitas yang digunakan. Kemudian kesulitan-kesulitan ataupun kekurangan yang terdapat pada penelitian NER dapat diketahui. Dapat dilihat

seperti pada penelitian [4] bahwa jika hanya menerapkan *word embedding* tanpa dikombinasikan *POS Tag*, kinerjanya masih kurang tetapi jika dikombinasikan maka akan meningkatkan akurasi sebanyak 13.12%. Kemudian pada [2] diketahui jika model yang lebih kompleks yaitu BiLSTM-CNN-LSTM justru mendapatkan skor terkecil. Menurut peneliti masalah tersebut terjadi karena *dataset*-nya yang masih sedikit sehingga tidak dapat mengakomodasi model kompleks untuk dipelajari. Pada penelitian [6] terdapat entitas nama lokasi yang unik dan tidak ada dalam data latih serta *Gazetteer* ternyata dapat membuat sistem tidak mengenali kelas entitas dengan tepat. Penelitian [16] mengalami masalah karena terbatasnya data dan biaya sehingga hasil yang didapatkan kurang maksimal. Masalah yang sama juga dialami oleh [10] karena tidak tersedianya data perjalanan dan wisata standar maka peneliti belum dapat menentukan keakuratannya.

TABLE IV. TABEL HASIL DAN ENTITAS

Referensi	Metode	Hasil	Entitas
[3]	<i>Multitask</i> BiRNN	Memperoleh akurasi skor F sebesar 93.31%	<i>Disease, symptom, treatment, test, dan disease group.</i>
[4]	LSTM, BiLSTM	Mendapatkan skor F1 sebesar 77.08%, dengan kombinasi POS Tag dan Word Embedding dapat meningkatkan skor F1 sebesar 13.12%	<i>Organization, person, location.</i>
[5]	LSTM, BiLSTM	Dengan BiLSTM-RNN dipadukan dengan CRF mendapat skor F1 86.55% pada data GM dan 73.99% pada data JNLPBA.	<i>DNA, RNA, protein, cell line, cell type.</i>
[2]	BiLSTM, CNN	Pada kombinasi BiLSTM-CNN mendapatkan skor F1 79.43%	<i>Person, organization, location, event.</i>
[6]	<i>Recurrent</i> CNN	Memperoleh skor F1 93.53%	<i>Loc, Gpe, Bld, Npl, Hwymse, Obj, Mse, Time, Date, Other.</i>
[7]	BiLSTM, CRF	BiLSTM dengan CRF mampu memperoleh	<i>person, organization, time, quantity location dan other.</i>

		akurasi sebesar 87,77%	
[8]	BiLSTM, CRF	Skor F1 yang didapat sebesar 0.855.	<i>Name, Location, Judicial Organization, Docket Number dan Crime Type</i>
[9]	BiLSTM, CNN	Mendapatkan akurasi skor F1 91.62% pada CoNLL 2003 dan 86.28% pada OntoNotes.	<i>Location, organization, person, dan miscellaneous</i>
[10]	CRF	F1 skor yang didapat sebesar 83%.	<i>Location, person, organization, money, percent, date dan time</i>
[11]	BERT	Diketahui untuk dengan spaCy untuk entitas LOC/ORG mendapat akurasi 150.97, FAC sebesar 77.42, dan LOC/ORG/FAC 91.75 pada data <i>test</i> . Sedangkan dengan BERT untuk entitas LOC/ORG mendapat akurasi F1 0.258, FAC 0.245 dan LOC/ORG/FAC 0.464. Model ini masih menghasilkan error 8%-25%	<i>Location, organization, facility</i>
[12]	YATSI, NBC, KNN	Mendapatkan akurasi 70.43% dan <i>F-measure</i> 69%.	<i>Nature, place, city, region, negative class.</i>
[13]	RNN, CNN, CRF	Mendapatkan skor F1 sebesar 83.60% untuk CRF dan 82.77% untuk CNN. Sedangkan skor F1 sebesar 85.94% untuk RNN sekaligus	<i>Problem, Treatment, Test.</i>

		menjadi skor F1 tertinggi.	
[14]	BiLSTM, CRF	Mendapatkan skor F1 sebesar 75.5% dengan menggabungkan <i>Softmax-POS</i> .	<i>Person, loc, ind, evt, fnb.</i>
[15]	CNN	Skor F1 yang diperoleh sebesar 71.01% dengan menggabungkan <i>skip-gram</i> dan <i>POS Tag</i>	<i>protein, DNA, RNA, cell_type dan cell_line.</i>
[16]	BERT, BiLSTM, CRF	Menghasilkan skor F1 sebesar 85.52% untuk BiLSTM-CRF, sedangkan untuk CRF mendapatkan sebesar 88.88% dan untuk gabungan BERT-BiLSTM-CRF (BBLC) sebesar 84.79%.	<i>Name, location, organization, time, thing.</i>

#### IV. KESIMPULAN

Dalam tinjauan literatur ini, terdapat 15 literatur yang didapatkan melalui *Google Scholar*, *Elsevire*, *arXiv* dan *Medwell* dalam rentang tahun 2015 hingga 2020. Literatur telah dikaji dan dapat diketahui bahwa perkembangan tren dalam penelitian NER sebelumnya, penelitian dilakukan menggunakan *deep learning* dengan beberapa metode seperti LSTM, BiLSTM, CNN dan dapat dikombinasikan satu sama lain. Penggunaan metode BiLSTM sangat populer dan paling banyak digunakan oleh peneliti karena dari 15 literatur BiLSTM digunakan dalam 8 literatur. BiLSTM dapat memahami dan mengambil perspektif dari ekstraksi informasi terdahulu dan ekstraksi informasi terdapat, sehingga proses pembelajaran akan semakin dalam yang berdampak pada model akan lebih memahami konteks pada ekstraksi informasi tersebut sehingga BiLSTM akan bermanfaat dalam melabeli entitas secara sekuensial apabila memiliki akses terhadap kedua informasi dari informasi sebelum dan sesudahnya. Kemudian penggunaan ekstraksi fitur juga banyak dilakukan, hasilnya sangat mempengaruhi kinerja metode dan dapat meningkatkan hasil dari metode. Selain itu, saat ini tren penerapan NER beberapa tahun terakhir diterapkan untuk biomedis, medis, berita, media sosial dan *tourism*. Berdasarkan tinjauan literatur maka penelitian NER dalam bidang *tourism* masih sedikit dilakukan. Metode dengan hasil terbaik di bidang *tourism* diperoleh dengan menggunakan CRF. Penerapan NER pada bidang *tourism* masih harus dikembangkan lagi penelitiannya. Pengembangan penelitian pada bidang *tourism* bisa dilakukan dengan memperbanyak penelitiannya di bidang tersebut menggunakan *deep learning* yang lebih

bervariasi, memperbanyak data yang digunakan dan dapat mengenali kelas entitas yang lebih luas.

#### REFERENSI

- [1] "TripAdvisor Network Effect and the Benefits of Total Engagement | TripAdvisor Insights." [Online]. Available: <https://www.tripadvisor.com/TripAdvisorInsights/w828>. [Accessed: 24-Jun-2020].
- [2] W. Gunawan, D. Suhartono, F. Purnomo, and A. Ongko, "Named-Entity Recognition for Indonesian Language using Bidirectional LSTM-CNNs," *Procedia Comput. Sci.*, vol. 135, pp. 425–432, 2018.
- [3] S. Chowdhury *et al.*, "A multitask bi-directional RNN model for named entity recognition on Chinese electronic medical records," *BMC Bioinformatics*, vol. 19, no. Suppl 17, 2018.
- [4] V. Rachman, S. Savitri, F. Augustianti, and R. Mahendra, "Named entity recognition on Indonesian Twitter posts using long short-term memory networks," *2017 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2017*, vol. 2018-Janua, pp. 228–232, 2017.
- [5] C. Lyu, B. Chen, Y. Ren, and D. Ji, "Long short-term memory RNN for biomedical named entity recognition," *BMC Bioinformatics*, vol. 18, no. 1, pp. 1–11, 2017.
- [6] F. N. Putra and C. Fatchah, "Klasifikasi jenis kejadian menggunakan kombinasi NeuroNER dan Recurrent Convolutional Neural Network pada data Twitter," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 4, no. 2, p. 81, 2018.
- [7] H. Permana and K. K. Purnamasari, "Named Entity Recognition Using Bidirectional Lstm-Crf Methods in Indonesian Text," *Procedia Comput. Sci.*, no. 112, 2019.
- [8] P. Tang, P. Yang, Y. Shi, Y. Zhou, F. Lin, and Y. Wang, "Recognizing Chinese judicial named entity using BiLSTM-CRF," *J. Phys. Conf. Ser.*, vol. 1592, no. 1, 2020.
- [9] J. P. C. Chiu and E. Nichols, "Named Entity Recognition with Bidirectional LSTM-CNNs," *Trans. Assoc. Comput. Linguist.*, vol. 4, no. 2003, pp. 357–370, 2016.
- [10] J. Vijay and R. Sridhar, "A Machine Learning Approach to Named Entity Recognition for the Travel and tourism Domain," *Asian J. Inf. Technol.*, vol. 15, no. 21, pp. 4309–4317, 2016.
- [11] C. Chantrapornchai and A. Tunsakul, "Information Extraction based on Named Entity for Tourism Corpus," *JCSSE 2019 - 16th Int. Jt. Conf. Comput. Sci. Softw. Eng. Knowl. Evol. Towar. Singul. Man-Machine Intell.*, pp. 187–192, 2019.
- [12] K. E. Saputro, S. S. Kusumawardani, and S. Fauziati, "Development of semi-supervised named entity recognition to discover new tourism places," *Proc. - 2016 2nd Int. Conf. Sci. Technol. ICST 2016*, pp. 124–128, 2016.
- [13] Y. Wu, M. Jiang, J. Xu, D. Zhi, and H. Xu, "Clinical Named Entity Recognition Using Deep Learning Models," *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2017, pp. 1812–1819, 2017.
- [14] D. Hoesen and A. Purwarianti, "Investigating Bi-LSTM and CRF with POS Tag Embedding for Indonesian Named Entity Tagger," *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*, pp. 35–38, 2019.
- [15] L. Yao, H. Liu, Y. Liu, X. Li, and M. W. Anwar, "Biomedical Named Entity Recognition based on Deep Neural Network," *Int. J. Hybrid Inf. Technol.*, vol. 8, no. 8, pp. 279–288, 2015.

- [16] L. Xue, H. Cao, F. Ye, and Y. Qin, "A method of chinese tourism named entity recognition based on bblc model," *Proc. - 2019 IEEE SmartWorld, Ubiquitous Intell. Comput. Adv. Trust. Comput. Scalable Comput. Commun. Internet People Smart City Innov. SmartWorld/UIC/ATC/SCALCOM/IOP/SCI 2019*, no. September 1995, pp. 1722–1727, 2019.