



**DETEKSI LOKASI SISWA SMP DI INSTAGRAM DENGAN METODE NAMED ENTITY RECOGNITION**

**Sulthan Aliff Secca Ramadhani<sup>1</sup>, Alifio Rahmanqa<sup>2</sup> dan Nur Aini Rakhmawati<sup>3</sup>**

Institut Teknologi Sepuluh Nopember Surabaya Jawa Timur, Indonesia  
sultanaiff@gmail.com, alifiorahmanqa12092000@gmail.com dan nur.aini@is.its.ac.id

Diterima:

**7 Juli 2021**

Direvisi:

**8 Juli 2021**

Disetujui:

**14 Juli 2021**

**Abstrak**

*Instagram* dapat dipergunakan oleh semua kalangan baik dari anak-anak hingga dewasa. Melalui jejaring sosial yang luas masyarakat dapat melakukan sebuah kegiatan seperti berbagi suatu aktivitas cerita, membagi foto dan membagi video. Untuk itu perlu diketahui bahwa di Indonesia terutama anak SMP banyak sekali menggunakan aplikasi *Instagram*. Dengan adanya metode *Named Entity Recognition* (NER) kemungkinan besar untuk mengetahui lokasi anak SMP sangatlah besar. Metode tersebut bisa mengetahui sebuah lokasi berdasarkan *caption Instagram* yang dibagikan. Penelitian ini bertujuan untuk mengetahui lokasi yang sering dikunjungi oleh siswa SMP melalui *caption Instagram* dengan metode NER. Metode yang digunakan dalam penelitian ini yaitu dimulai dengan melakukan studi literatur guna dalam pembuatan tinjauan pustaka dan pendahuluan yang didasarkan dari kutipan para ahli dan penelitian sebelumnya. Setelah dilakukan analisis dengan metode NER terhadap total 200 *caption*, didapatkan entitas "*location*" berjumlah 60 dari 300 entitas lainnya. Dari total 60 entitas "*location*", penyebutan lokasi yang paling banyak adalah Indonesia dan Surabaya. Namun dari hasil analisis dengan metode NER yang kami lakukan, masih kurang optimal. Dengan beberapa kata yang tidak sesuai dengan entitasnya dan diakibatkan model tersebut *overfitting* terhadap *data training*.

**Kata kunci** : *Siswa SMP; Instagram; Metode Named Entity Recognition*

**Abstract**

*Instagram can be used by all circles from children to adults. Through a wide social network, people can do activities such as sharing a story, sharing photos and sharing videos. Therefore, please note that in Indonesia, especially junior high school children, there are many Instagram applications. With the Named Entity Recognition (NER) method, it is very likely to know the location of junior high school children. The method can find out a location based on Instagram captions shared. This research aims to find out the locations frequented by junior high school students through Instagram captions with NER method. The method used in this study is to start by conducting literature studies in the creation of library reviews and introductions based on expert quotes and previous research. After analysis by NER method of a total of 200 captions, 60 "LOCATION" entities were obtained from 300 other entities. Out of a total of 60 "LOCATION" entities, the most mentioned locations are Indonesia and Surabaya. But from the results of the analysis with the NER method that we did, still less optimal. With a few words that do not correspond to the entity and resulted from the model overfitting the training data.*

**Keywords**: *Junior high school students; Instagram; Named Entity Recognition Method*

## PENDAHULUAN

Pesatnya perkembangan teknologi telah membawa manusia untuk memungkinkan cara baru dalam berinteraksi dan bersosialisasi (Baharun & Finori, 2019), salah satunya adalah inovasi pada teknologi komunikasi yaitu media sosial (Putri Clarissa, 2020). Sebagai sarana komunikasi yang interaktif dan efektif, media sosial memiliki peran dalam aktivitas interaksi antar masyarakat (Purbohastuti, 2017). Di balik munculnya berbagai macam bentuk pertemanan media sosial yang kian digemari jutaan penduduk di dunia (Anindhita, 2018) telah membawa pergeseran nilai-nilai sosial dalam kehidupan masyarakat, khususnya remaja (Syam, 2015). Media sosial hadir menjadi bagian dari proses perjalanan tumbuh dewasa para remaja.

Dengan pesatnya perkembangan *new media*, telah memberikan dampak terhadap situs-situs komunikasi, dari sekedar *email* dan *chatting* (Nainggolan et al., 2018), menjadi media sosial atau jejaring sosial seperti *Facebook*, *Twitter*, *Instagram*, *Path* dan *Snapchat* (Mahendra, 2017). Salah satu media sosial yang sedang banyak digunakan oleh pengguna gadget saat ini adalah *Instagram* (Fauziah et al., 2018). *Instagram* mendefinisikan sebagai situs jejaring sosial dengan layanan berbasis *web* (Wicaksono & Rummyeni, 2017) yang memungkinkan perorangan membangun profil umum atau semi-umum dalam satu sistem yang terbatas (Karina, 2020), menampilkan pengguna lainnya yang berkaitan dengan mereka.

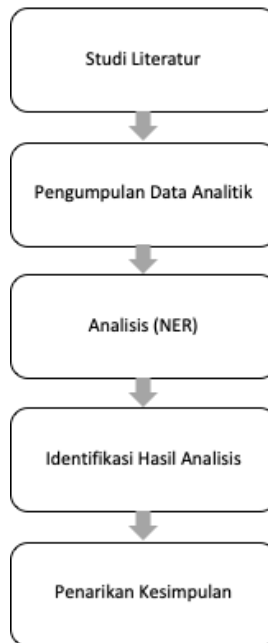
Media sosial *Instagram* telah membuat masyarakat luas untuk melakukan sebuah kebebasan berekspresi (Irawan & Yusuf, 2017). Tidak hanya orang dewasa saja yang dapat mengakses media sosial *Instagram* melainkan remaja hingga anak-anak (Zakirah, 2017) juga dapat mengakses media sosial tersebut. *Instagram* memungkinkan penggunanya untuk dapat mengunggah sebuah foto (Febrina, 2019) atau video untuk mengabadikan pengalaman dari sebuah kegiatan yang dilakukan (Aripradono, 2020). Selain itu, hal yang menarik adalah dapat membagikan sebuah postingan dengan memberikan titik keberadaan atau lokasi. Hal ini berfungsi agar pengguna dapat memberikan keterangan lokasi dari aktivitas yang dibagikan di *Instagram* (Roikanah & Rachman, 2021).

Pada dasarnya metode NER (*Named Entity Recognition*) digunakan untuk mengekstrak suatu informasi untuk tujuan didaptkannya suatu informasi yang terstruktur dan akurat dari teks yang tidak terstruktur. Pada saat itu, orang-orang menyadari bahwa sangat penting untuk mencari jenis informasi seperti nama orang, nama organisasi, nama tempat, waktu, tanggal, nominal uang dan persentase. Proses menandai jenis entitas dinamakan "*Named Entity Recognition*". Sehingga nantinya dengan metode ini, informasi-informasi dari berbagai *caption* postingan yang mengandung sebuah lokasi dapat terekstrak sehingga didapatkan informasi lokasi pada pengguna Instagram khususnya pelajar SMP. Penelitian ini bertujuan untuk untuk mengetahui lokasi yang sering dikunjungi oleh siswa SMP melalui *caption* instagram dengan metode NER. Manfaat penelitian yaitu mengetahui aktivitas penggunaan media sosial instagram pada siswa SMP.

## METODE PENELITIAN

Metode yang digunakan dalam penelitian ini yaitu dimulai dengan melakukan studi literatur guna dalam pembuatan tinjauan pustaka dan pendahuluan yang didasarkan dari kutipan para ahli dan penelitian sebelumnya. Selanjutnya, pengumpulan akun *Instagram* yang mengikuti akun resi SMP dan *caption* instagram yang mengandung sebuah lokasi,

kemudian dilanjutkan melakukan metode analisis dengan menggunakan metode NER . Setelah itu, dari hasil proses NER nantinya mengidentifikasi lokasi yang sering dikunjungi berdasarkan analisis dari kumpulan *caption* dari tiap akun instagram dan terakhir adalah penarikan kesimpulan dari hasil identifikasi.



Gambar 1. Metodologi Penelitian

Berdasarkan penelitian hal utama yang harus dilakukan adalah mengidentifikasi masalah. Dengan studi literatur kita dapat melakukan sebuah identifikasi masalah dengan tepat. Studi literatur pada penelitian kali ini mencakup pembahasan mengenai deteksi lokasi siswa SMP dengan penggunaan akun instagram berdasarkan *caption* yang telah dibuat menggunakan metode *Named Entity Recognition (NER)*. NER merupakan sebuah metode untuk mengetahui *information extraction* sebuah teks *caption* menjadi sebuah data. Tujuan metode tersebut mengidentifikasi sebuah kata menjadi dalam sebuah entitasnya. Dalam tahap ini, kami melakukan pengumpulan data yang diperlukan untuk tahap analisis. Data ini terdiri dari 20 akun *Instagram* seseorang yang mengikuti akun resmi SMP (Sekolah Menengah Pertama) serta 10 *caption* sebuah postingan instagram yang mengandung sebuah lokasi (berupa lokasi nama kota). Berikut merupakan data yang kami kumpulkan untuk nantinya dianalisis :

Tabel 1. Daftar Akun *Instagram* anak SMP beserta Sekolah SMP

No	Nama akun <i>Instagram</i>	Akun resmi SMP yang diikuti
1	__looking.caesario78__	smpmuhammadiyah9jakarta
2	Murritza__	smpkreatifannursurabayaput
3	Putra_k410ib	smpgiki3surabaya
4	nabilasafirau	smppercik
5	Hisyam_cahya	Smpi_alazhar14
6	Cellin_dp	smpn21semarang
7	Yuga_yuga_yuga	smpn21semarang
8	Faurellia_fadhilla	smppercik
9	Aurellia.putriastadewi	smpnegeri6sby
10	Adhiraja_kidz	smpnegeri6sby
11	davinnayaka	smpn2surabaya
12	rizkykumairahrizkykumairah	smpnegeri2mlg
13	arofix.spra__	smpnegeri14semarang
14	alinakrn02	smpnegeri2mlg
15	raisyaadhimaniar	Smpnegeri48bdg
16	chestaardiona86	smpnegeri48bdg
17	rnuhamad_raihan	smpnegeri48bdg
18	navarroghaizani	smpnegeri48bdg
19	Kaka_oficial_12	smpnegeri48bdg
20	zaskia_natasyaa	smpnegeri48bdg

Setelah nantinya data yang diperlukan sudah ditentukan, dalam tahap ini kami melakukan analisis *caption Instagram* dengan metode NER (*Named Entity Recognition*). Disini kami melakukan jenis metode NER SpaCy. Untuk tahapan pertama yang kami lakukan adalah melakukan pengelolaan data kami perlu *preprocessing* pada *caption Instagram* dengan menghilangkan beberapa simbol, *emoticon* ataupun tanda baca yang tidak dibutuhkan. Berikut adalah contoh hasil ekstraksi *caption* yang sudah dilakukan *preprocessing*.

Pandemi COVID-19 membatasi ruang gerak kita untuk melakukan kegiatan sekolah, namun tak menghalangi semangat saya untuk melaksanakan MPLS dari rumah. Saya [nama] bangga menjadi sisw[a/i] SMPN 48 Bandung kelas 7J dan siap melaksanakan MPLS dari rumah dengan Yakin,

Gambar 2. *Caption Pasca Preprocessing*

Pada langkah *preprocessing* ini, dilakukan guna untuk mendapatkan hasil yang maksimal. Dimana disini mengubah *caption* menjadi sebuah teks data yang sesuai dengan format SpaCy. Selanjutnya disini kami melakukan identifikasi dengan metode NER. Dimana sebelum dilakukan identifikasi, telah dilakukan *training* dengan menggunakan data pembelajaran model NER sebanyak 30 kali perulangan. Setelah pembelajaran selesai, dapat langsung dilakukan *testing* dengan *dataset* yang disesuaikan. Berikut adalah hasil *testing* yang telah dilakukan.

```
Love you all")
...: print("Entities", [(ent.text, ent.label) for ent in doc.ents])
Entities [('Hey', 'PERSON'), ('Surabaya', 'LOCATION'), ('Bonek', 'LOCATION'),
('Green', 'LOCATION'), ('Surabaya', 'LOCATION'), ('Bangun', 'PERSON'), ('Huhuhuu Mayan
Melihat', 'LOCATION'), ('Buah', 'ORGANIZATION'), ('Nama : Shakila', 'ORGANIZATION'),
('Indonesia', 'LOCATION'), ('Indonesia', 'LOCATION'), ('MARITZA', 'ORGANIZATION'),
('RAHARDI', 'ORGANIZATION'), ('BIASA', 'ORGANIZATION'), ('MARITZA', 'ORGANIZATION'),
('RAHARDI', 'ORGANIZATION'), ('BIASA', 'ORGANIZATION'), ('DIRGAHAYU', 'ORGANIZATION'),
('Indonesia', 'LOCATION'), ('Aktif', 'LOCATION'), ('Kreatif', 'LOCATION'),
('Inovatif', 'PERSON'), ('Menyebarkan', 'LOCATION'), ('Allahumma', 'PERSON'),
('BIASA', 'ORGANIZATION'), ('MARITZA', 'ORGANIZATION'), ('RAHARDI', 'ORGANIZATION'),
('Belajar', 'PERSON'), ('VII', 'ORGANIZATION'), ('Kreatif An-', 'PERSON'), ('VII',
'ORGANIZATION'), ('2', 'QUANTITY'), ('16 I', 'QUANTITY'), ('TGS', 'ORGANIZATION'),
('16 I', 'QUANTITY'), ('Video', 'PERSON'), ('Bus Jatim', 'LOCATION'), ('GK',
```

Gambar 3. Hasil *Testing* NER

Dalam tahap ini, kami melakukan identifikasi hasil analisis yang telah kami lakukan sebelumnya. Tahap ini kami mendefinisikan hasil analisis yang masih berupa kode dalam metode NER yang kami lakukan hingga menjadi sebuah informasi yang dapat digunakan dalam membuat sebuah kesimpulan nantinya. Tahap ini merupakan tahapan terakhir dari rangkaian proses yang telah dilakukan dari awal. Kami memberikan kesimpulan dari analisis yang telah kami lakukan. Kesimpulan ini nantinya membantu dalam penjelasan tujuan dan hasil analisis yang telah dilakukan.

### HASIL DAN PEMBAHASAN

Berdasarkan dari analisis yang telah dilakukan, 300 entitas yang terklasifikasi yang terdiri dari entitas “*Location*”, “*Organization*”, “*Person*”, “*Quality*” dan “*Time*”. Berdasarkan dari tabel 2, dapat diketahui bahwa entitas yang paling banyak adalah “*Organization*”.

Tabel 2. Entitas Hasil NER

No.	Entitas	Jumlah
1.	<i>Location</i>	60
2.	<i>Organization</i>	89
3.	<i>Person</i>	56
4.	<i>Quantity</i>	56
5.	<i>Time</i>	38

Dari total *caption* yang dikumpulkan mayoritas *caption* pada akun siswa SMP merupakan *caption* dari kegiatan seperti pengumpulan tugas sekolah, perkenalan identitas diri sebagai siswa baru, dan kata singkat yang mengandung sebuah ekspresi. Sedangkan, untuk *caption* yang mengandung lokasi bersumber dari penyebutan letak dari sekolahnya dan destinasi wisata. Berikut merupakan *detail* dari entitas “*Location*” :

Tabel 3. Entitas *Location*

No.	Lokasi	Jumlah
1.	Surabaya	5
2.	Bonek	1
3.	Green	1
4.	Indonesia	1
5.	Huhuhuu Mayan Mayan melihat	5
6.	Aktif	1
7.	Kreatif	1
8.	Menyebarkan	1
9.	Bus Jatim	1
10.	Kota Kediri Jawa Timur:Office	1

11.	Bersepeda	1
12.	Sleman	1
13.	Yogyakarta	1
14.	Uk.3	1
15.	Dapatkan	1
16.	Fattan	1
17.	B.Indonesia	1
18.	Larutan Asam	1
19.	Iseng	1
20.	Wassalamualaikum Partisipan 09	1
	Orang	1
21.	SMPN 6 Surabaya	1
22.	Ayooo	1
23.	Strategi	1
24.	Penyiraman	1
25.	Jembatan	1
26.	Penghematan	1
27.	Slide	1
28.	Wassalamualaikum	1
29.	Dufan	1
30.	Paris	1
31.	Winter	1
32.	Primary	1
33.	Maen	1
34.	Yui	1
35.	Teknik Sepakbola <i>Online Episode</i>	1
36.	Sunda poster nganggo basa sunda	1
37.	Ayeunna	1
38.	75 Guru	1
39.	Mantap	2
40.	Sediiiiih	1
41.	Braga Ciwalk Bandung	1
42.	Wong	1
43.	Adiwiyata	1
44.	Halo	1
45.	Busana	1
46.	Assalamualaikum wr wb	1
47.	Buat	2
48.	<i>New Normal</i>	1

Berdasarkan dari tabel 3, kata Indonesia dan Surabaya memiliki jumlah paling banyak dengan jumlah 5. Untuk kata yang lainnya hanya berjumlah 1 pada klasifikasi “*Location*” yang telah dilakukan. Hal ini disebabkan beberapa kata yang tidak sesuai karena model tersebut *overfitting* terhadap *data training*.

### **KESIMPULAN**

Berdasarkan penelitian yang telah dilakukan, disini kami ingin mengetahui lokasi yang paling sering dikunjungi oleh siswa SMP. Dimana disini mengambil dari 20 akun siswa yang mengikuti akun SMP dan 10 *caption* dari masing akun yang mengandung sebuah lokasi. Setelah dilakukan analisis dengan metode NER terhadap total 200 *caption*, didapatkan entitas “*Location*” berjumlah 60 dari 300 entitas lainnya. Dari total 60 entitas

“Location”, penyebutan lokasi yang paling banyak adalah Indonesia dan Surabaya. Namun dari hasil analisis dengan metode NER yang kami lakukan, masih kurang optimal. Dengan beberapa kata yang tidak sesuai dengan entitasnya dan diakibatkan model tersebut *overfitting* terhadap *data training*.

## BIBLIOGRAFI

- Anindhita, W. (2018). Dramaturgi Dibalik Kehidupan Sosial Climber. *Jurnal Komunikasi Dan Bisnis*, 6(1).
- Aripradono, H. W. (2020). Penerapan Komunikasi Digital Storytelling Pada Media Sosial Instagram. *Teknika*, 9(2), 121–128.
- Baharun, H., & Finori, F. D. (2019). Smart Techno Parenting: Alternatif Pendidikan Anak Pada Era Teknologi Digital. *Jurnal Tatsqif*, 17(1), 52–69.
- Fauziah, R., Ratnamulyani, I. A., & Kusumadinata, A. A. (2018). Efektifitas Promosi Destinasi Wisata Rekreasi Gunung Pancar Melalui Postingan Instagram Media Sosial. *Jurnal Komunikatio*, 4(1).
- Febrina, A. (2019). Motif Orang Tua Mengunggah Foto Anak Di Instagram (Studi Fenomenologi Terhadap Orang Tua di Jabodetabek). *Jurnal Abdi Ilmu*, 12(1), 55–65.
- Irawan, E., & Yusuf, Y. (2017). *Instagram Sebagai Gaya Hidup Masyarakat Kota Pekanbaru (Studi Komunitas Instagram Di Kota Pekanbaru)*. Riau University.
- Karina, M. E. (2020). Analisis Personal Branding Selebgram di Sosial Media Instagram (Studi Deskriptif Kualitatif Pada Akun Instagram@dindamaulinaaaa). *Jurnal Humas & Media Kontemporer*, 1(1), 11–24.
- Mahendra, B. (2017). Eksistensi sosial remaja dalam Instagram (sebuah perspektif komunikasi). *Jurnal Visi Komunikasi*, 16(1), 151–160.
- Nainggolan, V., Rondonuwu, S. A., & Waleleng, G. J. (2018). Peranan media sosial Instagram dalam interaksi sosial antar mahasiswa fakultas ilmu sosial dan politik UNSRAT Manado. *ACTA DIURNA KOMUNIKASI*, 7(4).
- Purbohastuti, A. W. (2017). Efektivitas Media Sosial Sebagai Media Promosi. *Tirtayasa Ekonomika*, 12(2), 212–231.
- Putri Clarissa, N. (2020). *Proses Pengelolaan Media Sosial Instagram Badan Litbang dan Inovasi KLHK Tahun 2019*. IPB University.
- Roikanah, U., & Rachman, M. A. (2021). Penggunaan Instagram Sebagai Sarana Pemasaran: Studi Kasus Taman Baca Masyarakat Kolong Ciputat. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi Dan Kearsipan*, 7(1), 1–18.
- Syam, H. M. (2015). Globalisasi Media dan Penyerapan Budaya Asing, Analisis Pada Pengaruh Budaya Populerkorea Di Kalangan Remaja Kota Banda Aceh. *Avant Garde*, 3(1).
- Wicaksono, M. A., & Rummyeni, R. (2017). *Pengaruh media sosial instagram@ wisatadakhokura terhadap minat berkunjung followers*. Riau University.
- Zakirah, D. M. A. (2017). *Mahasiswa dan Instagram (Study Tentang Instagram Sebagai Sarana Membentuk Citra Diri di Kalangan Mahasiswa Universitas Airlangga)*. Universitas Airlangga.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License