

Post-Edit Generation terhadap Analisa Wisata dan Kuliner

Herdian Pratama¹, Riska Aryanti², Rahmat Hidayat³

¹Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri/Illmu Komputer
e-mail: herdianpratama76@gmail.com

²universitas Bina Sarana Informatika/Illmu Komputer
e-mail: riska.rts@bsi.ac.id

³universitas Bina Sarana Informatika/Sistem Informasi
e-mail: rahmat.rhh@bsi.ac.id

Abstrak

Purwakarta merupakan salah satu daerah di Jawa Barat yang masyarakatnya tumbuh dan berkembang dari sejarah dan tradisi yang kental. Purwakarta memiliki banyak objek wisata yang dikelola oleh pemerintah maupun pihak swasta, selain objek wisata purwakarta juga dikenal dengan kulinernya yang sangat khas dengan cita rasa yang tradisional, tidak sedikit wisatawan purwakarta yang merasa bingung dalam memutuskan untuk mengunjungi tempat wisata yang cocok dikarenakan kurangnya informasi mengenai objek wisata maupun kuliner tersebut. Untuk itu dibutuhkan suatu sistem yang dapat membantu wisatawan untuk mendapatkan ringkasan informasi mengenai wisata dan kuliner terdekat serta memberikan rekomendasi sehingga peningkatan efisiensi dan efektifitas dalam pendistribusian informasi untuk calon wisatawan yang ingin berkunjung dan laporan hasil akhirnya terperinci mengenai kumpulan objek wisata dan kuliner yang ada di purwakarta sehingga memudahkan pengguna, yang mana inputnya berdasarkan deskripsi (Nama Tempat, Alamat, Fasilitas, Wahana) serta ulasan pengguna lainnya. Untuk itu penelitian ini menyajikan studi kasus Post-Edit Generation terhadap wisata dan kuliner di Purwakarta-Jawa Barat.

Kata Kunci: Post-Edit Generation, Matriks Evaluasi, Wisata dan Kuliner

Abstract

Purwakarta is one of the areas in West Java where the community grows and develops from thick history and tradition. Purwakarta has many tourist objects managed by the government and private parties, in addition to Purwakarta tourist attractions are also known for their culinary that is very distinctive with traditional flavors, not a few Purwakarta tourists who feel confused in deciding to visit suitable tourist attractions due to lack of information about the object tourism and culinary. For that we need a system that can help tourists to get a summary of information about tourism and the nearest culinary and provide recommendations so as to increase efficiency and effectiveness in the distribution of information for prospective tourists who want to visit and detailed results report on the collection of attractions and culinary in Purwakarta. making it easier for users, which input is based on the description (Place Name, Address, Facilities, Forum) and other user reviews. For this reason, this research presents a Post-Edit Generation case study of tourism and culinary in Purwakarta-West Java.

Keywords: Post-Edit Generation, Matrix Evaluation, Tourism and Culinary

1. Pendahuluan

Wisata merupakan kebutuhan seseorang untuk menenangkan pikiran dari kepenatan sehari-hari, berwisata memang merupakan suatu hal yang amat didambakan, karena dengan berwisata

dapat sejenak melupakan keadaan sehari-hari yang monoton. Meskipun terkadang membutuhkan dana, namun tidak akan pernah sebanding dengan kepuasan rohani dan jasmani yang didapatkan. Menurut menteri pariwisata menyatakan bahwa kunjungan wisatawan ke purwakarta pada

November 2016 sebanyak 903.885 yang terdiri dari (1.280 Wisatawan Mancanegara dan 902.605 Wisatawan Nusantara) mengalami kenaikan yang signifikan dibandingkan tahun-tahun sebelumnya. Wisata purwakarta merupakan salah satu daerah yang mempunyai objek wisata yang menarik. Semakin berkembangnya dunia pariwisata khususnya di daerah purwakarta, namun tidak diiringi dengan kemajuan teknologi yang digunakan. Ditambah lagi dengan kondisi banyaknya wisatawan yang tidak mau berwisata karena merasa mininnya informasi mengenai objek wisata maupun kuliner di daerah tersebut. Sehingga keadaan ini membuat calon wisatawan yang ingin menghabiskan liburannya untuk berwisata kurang tepat dalam memutuskan untuk mengunjungi tempat wisata yang cocok untuk dikunjungi. Untuk itu dibutuhkan suatu sistem yang mampu memberikan ringkasan informasi yang memberikan laporan terperinci mengenai pariwisata maupun kuliner yang ada di daerah purwakarta.

Text edit generation sering membutuhkan pasca editing secara manual, apabila pengeditan dilakukan dengan benar maka akan menghasilkan data yang berkualitas tinggi (Han et al., 2019). Beberapa penelitian sebelumnya juga menunjukkan bahwa post-edit bisa dilakukan secara otomatis, pada penelitian Wilker Aziz ini membandingkan *machine translation* yang mana membandingkan empat *machine translation*, yakni Google, Musa, Systran dan Trados dengan tujuan mengevaluasi kualitas terjemahan (Aziz, De Sousa, & Specia, 2012), adapun pada penelitian Bojar fokus pada perbaikan substansial dalam mesin penerjemahan (Berita, IT-Domain, Biomedis) dengan menggunakan matrik evaluasi Bleu dan Hter (Bojar, Ondřej; Chatterjee et al., 2016).

Penelitian yang telah dilakukan menurut Carolina Scarton mengevaluasi kualitas hasil akhir dari *machine translation* dan *Automatic Summarisation* pada dokumen-level dengan tujuan untuk mengeksplorasi tingkatan dokumen, dengan fokus pada fitur wacananya, namun matrik evaluasi untuk MT (Machine Translation) dan Automatic Summarisation seharusnya mampu mengukur kualitas dengan aspek yang berbeda (Scarton, 2015). Selain itu evaluasi post-edit juga dapat digunakan untuk menghasilkan

prakiraan cuaca dengan SumTime-Mousam, pada penelitian ini menggunakan post-edit untuk memperbaiki sistem, tidak hanya untuk mengukur kinerjanya (Sripada, Reiter, & Hawizy, 2005).

Penelitian yang telah dilakukan oleh Julie Carson penerapan Post Edit untuk *speech recognition* dan analisa text (Berndsen-Carson, Somers, & Way, 2010). Sedangkan pada penelitian Randy Scansani penerapan post-edit untuk mesin translate Bahasa Italia ke Bahasa Inggris (Scansani, Ferraresi, Bernardini, & Bentivogli, 2019), *Post-Edit Generation* juga digunakan untuk pembuatan biografi manusia dengan evaluasi post edit dari 525.555 token yang disunting oleh sistem dengan matrik evaluasi menggunakan Rouge dan Bleu dengan tujuan meringkasnya informasi tentang seseorang di web (Han et al., 2019), namun pada penelitian ini kurang jelasnya aturan editorial dimana si pengguna mengalami masalah dalam pengeditan dan kurangnya kebijakan editorial yang jelas dengan pengguna terkadang membuat suntingan yang saling bertentangan.

Adapun tujuan riset dalam penelitian ini adalah membuat suatu sistem yang membantu pengguna dapat dengan mudah mendapatkan ringkasan informasi tentang wisata beserta kuliner di daerah tersebut, membantu pengguna untuk mengetahui popularitas objek wisata beserta kulinernya pada suatu tempat tersebut, mendapatkan rekomendasi tempat wisata beserta kulinernya pada daerah tersebut. Untuk itu peneliti menerapkan perubahan yang disarankan di atas, dengan inovasi dari penelitian sebelumnya.

Ide dasar paper tersebutlah yang mendasari pengembangan riset pada penelitian ini. Namun selanjutnya muncul pertanyaan ketika pengguna ingin menggunakan post-edit untuk menghasilkan informasi ringkasan yang terperinci dari pariwisata dan kuliner yang ada di purwakarta, antara lain:

1. Bagaimana post-edit dapat diterapkan untuk mengetahui ringkasan yang terperinci mengenai pariwisata dan kuliner yang ada di purwakarta?
2. Apakah pengguna dapat dengan mudah mengetahui informasi tentang wisata maupun kuliner yang ada di daerah purwakarta, dengan ringkasan informasi secara terperinci tidak hanya

dari ulasan saja namun deskripsi(Nama Tempat, Alamat, Fasilitas, Wahana) nya juga?

3. Bagaimana pengguna dapat mengetahui popularitas objek wisata beserta kulinernya pada suatu tempat tersebut?

Dari ketiga pertanyaan tersebut belum mampu terjawab dengan riset di atas. Hal inilah yang menjadikan kelemahan pada paper tersebut yang selanjutnya perlu diatasi. Maka dari itu peneliti mengambil judul “*Post-Edit Generation* terhadap Analisa Wisata dan Kuliner”.

2. Metode Penelitian

Metodologi yang akan digunakan dalam pengembangan *Post-Edit Generation* pada wisata dan kuliner berdasarkan deskripsi dan ulasan adalah Design Research karena akan membuat suatu *Artifak* berupa aplikasi. Sedangkan teknik pencarian solusi yang digunakan adalah teknik *modeling existing solution* Adapun tahapannya yaitu:

1. Dataset yang digunakan untuk mendapatkan kumpulan informasi wisata dan kuliner yang ada didapatkan dari situs-situs web, berita maupun sosial media, seperti twitter, facebook, instagram.
2. Untuk pengeditan text pada penelitian sebelumnya menggunakan *diff library phyton*, namun kali ini pada penelitian ini menggunakan *java diff utils library*
3. Dalam Menentukan Karakteristik Pengeditan digunakan dengan menggunakan metode penelitian dari sebelumnya(Han et al., 2019), yang mana

dari kalimat yang ada dibagi menjadi beberapa kategori sebagai berikut :

- a. Kata yang dipilih
Kalimat yang dipilih adalah kalimat yang berisi Deskripsi(Nama Tempat, Alamat, Fasilitas, Wahana,.) serta ulasan pengguna yang dianggap penting.
Dari kalimat yang dipilih kemudian kalimat dibagi menjadi tiga kategori, yang mana:

- 1) Kata yang diedit
Pemilihan dari token pada kalimat yang umumnya bersifat *gramatikal* (pengeditan hanya membahas

konsistensi atau keselarasan kapitalisasi pada pengumpulan deskripsi maupun ulasan)

- 2) Kata yang dihapus
Terdapat beberapa kalimat yang akan dihapus, dikarenakan kalimat yang mengandung perjalanan seseorang dan kalimat yang tumpang tindih dengan kalimat hasil akhir.
 - 3) Kata Baru yang disisipkan
Menambahkan kalimat baru yang mana berisi deskripsi atau ulasan dari wisata maupun kuliner yang lebih bermanfaat. Kalimat baru sebagian besar termasuk konten wisata maupun kuliner yang disarankan namun kalimat tersebut masih merupakan kalimat yang signifikan untuk meningkatkan hasil akhir dari ringkasan informasi mengenai wisata dan kuliner di purwakarta.
 - 4) Kata yang tidak dipilih
Kalimat yang tumpang tindih dengan kalimat hasil akhir pengeditan.
4. Menentukan Visualisasi Edit-flow(Han et al., 2019).

- a. *Sentence alignment* yakni penyelarasan kalimat yang disarankan dengan kalimat

terakhir dengan menggunakan *lavenshtein ratio*.

Contoh :

Misalnya, jarak *Levenshtein* antara "lavenshtein" dan "meilenshtein" adalah 4, karena empat pengeditan berikut mengubah satu ke yang lain, dan tidak ada cara untuk melakukannya dengan kurang dari empat suntingan:

		m	e	i	l	e	n	s	t	e	i	n
l	0	1	2	3	4	5	6	7	8	9	10	11
e	1	1	2	3	3	4	5	6	7	8	9	10
v	2	2	1	2	3	3	4	5	6	7	8	9
e	3	3	2	2	3	4	4	5	6	7	8	9
n	4	4	3	3	3	3	4	5	6	6	7	8
s	5	5	4	4	4	4	3	4	5	6	7	7
h	6	6	5	5	5	5	4	3	4	5	6	7
t	7	7	6	6	6	6	5	4	4	5	6	7
e	8	8	7	7	7	7	6	5	4	5	6	7
i	9	9	8	8	8	7	7	6	5	4	5	6
n	10	10	9	8	9	8	8	7	6	5	4	5
	11	11	10	9	9	9	8	8	7	6	5	4

```

l e v e n s h t e i n
o = o + = = = - = = = =
m e i l e n s t e i n
    
```

atau

l	e	v	e	n	s	h	t	e	i	n
o	=	+	o	=	=	=	.	=	=	=
m	e	i	e	n	s	t	e	i	n	

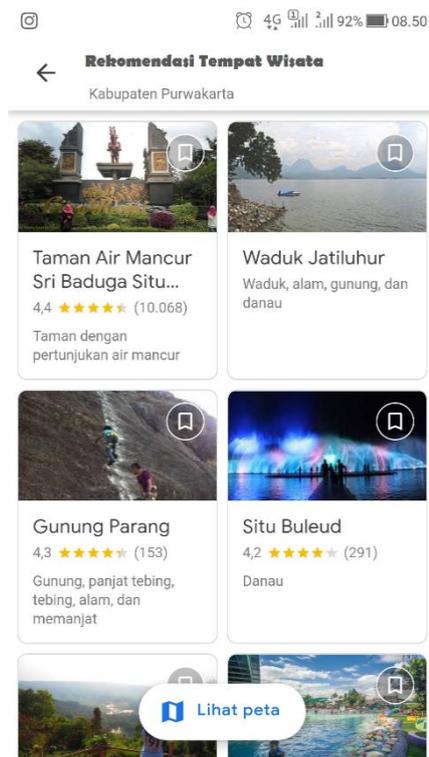
"=" Match; "o" Substitution; "+" Insertion; "." Deletion

- b. *Edit Flow*, yakni jumlah keseluruhan kalimat yang disarankan dibagi menjadi 2 kategori yang mana terdiri dari kalimat yang dipilih dan kalimat yang tidak dipilih berdasarkan karakteristik pengeditan
 - c. *Token Aligment*, memperbaiki frase pengganti dari kalimat yang selaras dengan pengeditan menggunakan java diff utils library.
5. Matrik evaluasi yang digunakan untuk penentu keputusan pada penelitian ini menggunakan Rouge dan Bleu dengan menambahkan matrik evaluasi Hter untuk penentuan rata-rata sehingga meningkatkan hasil akhir dari post-edit wisata dan kuliner dipurwakarta.
 6. Hasil dari ringkasan tersebut diintegrasikan dengan google map, sehingga mampu memberikan rekomendasi tempat wisata maupun kuliner terdekat yang ada di purwakarta bukan hanya dari ulasanya saja melainkan deskripsi informasi dari tempat wisata maupun kulinernya juga.

Pada tahapan selanjutnya sistem post-edit yang terbentuk dengan integrasi dengan google map akan diuji oleh wisatawan yang akan berwisata kepurwakarta dan akan dievaluasi bagaimana sistem tersebut akan mampu memberikan informasi yang terperinci mengenai wisata maupun kuliner yang ada di daerah tersebut, kemudian diidentifikasi kembali apakah sudah secara tepat memberikan rekomendasi kepada pengguna mengenai ketepatan informasi wisata maupun kuliner yang ada dipurwakarta.

3. Hasil dan Pembahasan

Berikut adalah tampilan hasil sistem penerapan *Post-Edit Generation* terhadap wisata dan kuliner yang ada dipurwakarta.



Gambar 2. Hasil Rekomendasi Tempat Pariwisata Terdekat



Gambar 3. Hasil Ringkasan Tempat Pariwisata



Ulasan

Anonymous
30 ulasan

★★★★★ 13 jam yang lalu

Tamannya luas dan terawat. Kolamnya di isi dengan ikan-ikan.

Gambar 4. Hasil Ulasan Tempat Pariwisata



Gambar 5. Hasil Foto Tempat Pariwisata



Gambar 6. Hasil Pencarian Kulinier Terdekat



Gambar 7. Hasil Ringkasan Kuliner



Berikan nilai dan ulasan
Bagikan pengalaman Anda untuk membantu orang lain



Ulasan

Anonymous
3 ulasan

★★★★★ 20 jam yang lalu

Tempatnya cozy... ayam gorengnya dan babat

Gambar 8. Hasil Ulasan Kuliner



Gambar 9. Hasil Foto Kuliner

4. Kesimpulan

Kesimpulan yang didapat pada pemaparan diatas bahwa penerapan *Post-Edit Generation* terhadap analisa wisata dan kuliner dapat membantu wisatawan untuk mendapatkan ringkasan informasi yang laporan hasil akhirnya terperinci mengenai kumpulan objek wisata dan kuliner yang ada di purwakarta sehingga memudahkan pengguna untuk mendapatkan informasi tentang objek wisata maupun kuliner yang ada di purwakarta, yang mana inputnya berdasarkan deskripsi (Nama Tempat, Alamat, Fasilitas, Wahana,.) serta ulasan pengguna lainnya.

Memberikan kemudahan kepada wisatawan untuk mendapatkan informasi mengenai wisata dan kuliner terdekat serta memberikan rekomendasi sehingga peningkatan efisiensi dan efektifitas dalam pendistribusian informasi untuk calon wisatawan yang ingin berkunjung ke purwakarta.

Referensi

- Aziz, W., De Sousa, S. C. M., & Specia, L. (2012). PET: A Tool for Post-Editing and Assessing Machine Translation. *Proceedings of the 8th International Conference on Language Resources and Evaluation, LREC 2012*, 3982–3987.
- Berndsen-Carson, J., Somers, H., & Way, A. (2010). Integrated language technology as a part of next generation localization [in] *Localisation Focus*. *Localisation Focus*, 8(1), 53–66. Retrieved from <http://www.localisation.ie/locfocus/issues/8/1>
- Bojar, Ondřej; Chatterjee, R. et al., Bojar, O., Chatterjee, R., Federmann, C., Graham, Y., Haddow, B., ... Zampieri, M. (2016). Findings of the 2016 Conference on Machine Translation (WMT16). *Proceedings of the First Conference on Machine Translation*, 2, 131–198. Retrieved from <http://www.aclweb.org/anthology/W/W16/W16-2301>
- Han, B., Radford, W., Cadilhac, A., Harol, A., Chisholm, A., & Hachey, B. (2019). Post-edit Analysis of Collective

-
- Biography Generation. *26th International World Wide Web Conference 2017, WWW 2017 Companion*, 791–792. <https://doi.org/10.1145/3041021.3054264>
- Scansani, R., Ferraresi, A., Bernardini, S., & Bentivogli, L. (2019). *Do translator trainees trust machine translation? An experiment on post-editing and revision*. 2, 73–79.
- Scarton, C. (2015). *Discourse and Document-level Information for Evaluating Language Output Tasks*. 118–125. <https://doi.org/10.3115/v1/n15-2016>
- Sripada, S. G., Reiter, E., & Hawizy, L. (2005). Evaluating an NLG System using Post-Editing. *IJCAI International Joint Conference on Artificial Intelligence*, 1700–1701.