

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Pekerja lepas atau *freelancer* kini menjadi pilihan populer. *Freelancer* adalah individu yang bekerja dengan waktu yang ditentukan sendiri tanpa terikat oleh perjanjian kerja. *Freelancer* dapat mengerjakan berbagai proyek seperti pemrograman, desain grafis, pengembangan situs web, dan penulisan artikel[1]. Pengaruh teknologi digital dalam era globalisasi ini telah merubah pandangan masyarakat tentang pekerjaan *freelancer*. Hal ini terlihat dari jumlah *freelancer* yang semakin bertambah. Menurut survei Badan Pusat Statistik (BPS) pada Agustus 2018, sekitar 56,8% dari total pekerja adalah *freelancer*, dengan sekitar 5,89 juta orang yang bekerja sebagai *freelancer*[2]. Platform berbasis *website* yang menyediakan layanan pencarian dan rekomendasi *freelancer* menjadi sangat populer, seperti Fastwork, Sribu, *Freelancer*, dan platform *freelancer* lainnya. Namun, dengan meningkatnya jumlah *freelancer* dan variasi keahlian yang ditawarkan, tantangan dalam menemukan *freelancer* yang tepat sesuai dengan kebutuhan spesifik klien juga semakin besar. Karena meningkatnya *freelancer* dan penawaran jasa yang tersedia, maka memerlukan metode yang efektif untuk merekomendasikan *freelancer* yang sesuai dengan yang dibutuhkan oleh pengguna.

Masalah utama yang dihadapi pengguna adalah kesulitan dalam mengidentifikasi *freelancer* yang paling cocok berdasarkan kriteria spesifik

yang diinginkan. Banyak pengguna yang merasa kesulitan dalam mengartikulasikan kebutuhannya dalam bentuk kata kunci yang efektif, yang sering kali menyebabkan hasil pencarian yang tidak relevan. Selain itu, sistem pencarian dan rekomendasi pada *platform* yang ada sering kali masih menggunakan metode pencocokan kata kunci yang sederhana, yang tidak mampu menangkap kompleksitas kebutuhan pengguna dengan akurat. Dengan menggunakan model *deep learning* seperti LSTM (*Long Shot Term Memory*) dan metode *cosine similarity*, dapat dilakukan analisis teks yang canggih untuk memahami dan membandingkan kesesuaian antara kebutuhan pengguna dan profil *freelancer*.

Namun, belum banyak penelitian yang secara khusus menerapkan model LSTM dan *cosine similarity* dalam konteks rekomendasi *freelancer* di *platform* jasa berbasis *website*. Oleh karena itu, penelitian ini akan menjadi kontribusi baru dalam pengembangan sistem rekomendasi dengan memanfaatkan model yang dapat digunakan secara efektif dalam konteks rekomendasi *freelancer*.

Untuk menyelesaikan masalah tersebut, penelitian ini akan mengumpulkan data profil *freelancer* dari *platform* layanan jasa yang relevan dan data kebutuhan pengguna dari *platform* media sosial. Selanjutnya, model LSTM akan digunakan untuk menghasilkan kata kunci dari kebutuhan pengguna, sedangkan *cosine similarity* akan digunakan untuk menghitung kesamaan antara kata kunci dari kebutuhan pengguna dengan data profil *freelancer*. Dengan mengintegrasikan kedua metode ini, diharapkan dapat

memberikan rekomendasi *freelancer* yang lebih akurat dan relevan bagi pengguna.

## 1.2. Tujuan dan Manfaat

### 1.2.1. Tujuan

- Mengembangkan dan mengoptimalkan model *Long Short-Term Memory* (LSTM) untuk memahami kebutuhan pengguna.
- Menerapkan metode *cosine similarity* untuk mencocokkan kebutuhan pengguna dengan profil *freelancer*
- Mengembangkan sistem rekomendasi menggunakan model LSTM dan metode *cosine similarity* untuk merekomendasikan *freelancer* kepada pengguna.

### 1.2.2. Manfaat

- Menyediakan metode baru dalam pencarian *freelancer* yang memanfaatkan model LSTM dan *cosine similarity*, yang dapat dijadikan referensi bagi penelitian selanjutnya.
- Mendorong pengembangan dalam mengimplementasikan model LSTM dan metode *cosine similarity* dalam proyek nyata.
- Membantu pengguna platform pencarian *freelancer* menemukan *freelancer* yang sesuai dengan kebutuhan pengguna.

### 1.3. Tinjauan Pustaka

Tinjauan pustaka ini bertujuan untuk memberikan gambaran mengenai penelitian terdahulu yang relevan dengan topik penerapan model *Long Short-Term Memory* (LSTM) dan *cosine similarity*. Berikut adalah ulasan dari beberapa penelitian yang relevan dalam lima tahun terakhir:

Penelitian yang berjudul "Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM" membahas penggunaan LSTM untuk analisis sentimen terhadap kebijakan MBKM di Twitter. Penelitian ini menunjukkan bahwa LSTM efektif dalam mengklasifikasikan emosi dari *tweet* ke dalam enam kategori, dengan akurasi mencapai 80,42%. Temuan ini menyoroti dominasi sentimen "bingung" di antara pengguna, yang penting untuk evaluasi dan perbaikan kebijakan MBKM[3].

Dalam studi yang berjudul "Uji Kemiripan Kalimat Judul Tugas Akhir dengan Metode *Cosine similarity* dan Pembobotan TF-IDF" mengkaji penggunaan metode *cosine similarity* dan pembobotan TF-IDF untuk mengevaluasi kemiripan judul tugas akhir mahasiswa di Deli Husada *Delitua Health Institute*. Penelitian ini menunjukkan bahwa metode tersebut efektif dalam mengidentifikasi kemiripan judul, dengan hasil bahwa 43% judul tidak layak untuk diajukan kembali, sedangkan 53% dianggap layak karena tingkat kemiripan yang tinggi. Penelitian ini menyoroti pentingnya penggunaan teknik berbasis teks untuk mengurangi kemiripan dan meningkatkan keunikan judul tugas akhir mahasiswa[4].

Penelitian yang berjudul "Prediksi Genre Film Dengan Klasifikasi Multi Kelas Sinopsis Menggunakan Jaringan LSTM" membahas penggunaan LSTM untuk memprediksi genre film berdasarkan sinopsis. Penelitian ini menggunakan *dataset* dari *Kaggle* yang berisi sinopsis film dengan berbagai genre. Sinopsis ini diolah melalui proses tokenisasi, vektorisasi, dan *embedding* kata menggunakan NLP. Hasilnya, metode klasifikasi multi kelas dengan jaringan LSTM mencapai akurasi 98% dan *loss* 5%, menunjukkan efektivitas LSTM dalam memprediksi genre film berdasarkan sinopsis. Temuan ini memberikan dasar kuat untuk pengembangan sistem rekomendasi otomatis di industri perfilman[5].

Penelitian berjudul "Meningkatkan Akurasi *Long-Short Term Memory* (LSTM) pada Analisis Sentimen Vaksin Covid-19 di Twitter dengan *GloVe*" membahas optimalisasi model LSTM untuk analisis sentimen di Twitter terkait vaksin Covid-19. Dengan munculnya Covid-19 pada awal 2020, banyak diskusi tentang pengadaan vaksin terjadi di Twitter, sehingga diperlukan analisis sentimen terhadap opini publik. Penelitian ini menggunakan LSTM, yang awalnya kurang akurat dibandingkan dengan metode lain seperti Bi-LSTM. Untuk meningkatkan akurasi, digunakan metode *GloVe*, yang mengonversi kemunculan kata menjadi vektor untuk digunakan dalam lapisan *embedding* LSTM. Hasil penelitian menunjukkan bahwa penerapan *GloVe* meningkatkan akurasi LSTM dari 87% menjadi 89%, mendekati akurasi Bi-LSTM sebesar 88%. Ini menunjukkan bahwa metode *GloVe* dapat meningkatkan kinerja model LSTM dalam analisis sentimen[6].

Penelitian yang berjudul "Penerapan Algoritma *Cosine Similarity* dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta" membahas penerapan teknologi informasi dalam mengelola pertanyaan yang sering diajukan (FAQ) selama proses Penerimaan Mahasiswa Baru (PMB) di perguruan tinggi. Dalam penelitian ini, pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*) digunakan untuk memberikan bobot pada kata-kata dalam dokumen berdasarkan frekuensi kemunculannya dan inversi frekuensi dokumen yang mengandung kata tersebut. Algoritma *cosine similarity* digunakan untuk menghitung tingkat kesamaan antara dua objek yang dinyatakan dalam bentuk vektor kata kunci. Penelitian ini menggunakan 7 sampel data FAQ yang diperoleh melalui wawancara dengan Ibu Susilawati, S.Kom. Data tersebut kemudian diproses melalui tahapan *preprocessing*, pembobotan TF-IDF, dan *cosine similarity* untuk menentukan kesamaan tertinggi sebagai hasil akhir. Hasil penelitian menunjukkan bahwa penggunaan metode ini dapat mencapai tingkat akurasi hingga 64.28%, menunjukkan efektivitas pembobotan TF-IDF dan *cosine similarity* dalam sistem penjawab FAQ untuk PMB[7].

Penelitian berjudul "Penerapan Metode LSTM dalam Pembuatan Sistem Pendeteksi Berita Palsu Berbahasa Indonesia" membahas penggunaan teknologi LSTM untuk mendeteksi berita palsu. Dengan perkembangan teknologi, masyarakat semakin mudah mengakses berita melalui internet dan media sosial, namun sering kali berita yang ditemukan adalah *hoaks*. Untuk mengatasi hal ini, penelitian ini menggunakan LSTM dan *Indobert* untuk

membedakan berita palsu dari berita yang benar. *Dataset* yang digunakan terdiri dari 3309 berita *hoaks* dan *non-hoaks*, dan hasilnya menunjukkan bahwa model ini memiliki akurasi sebesar 94.2%, *precision* 0.944, *recall* 0.940, dan *f1-score* 0.941. Hasil penelitian ini menunjukkan bahwa LSTM efektif untuk mendeteksi berita palsu berbahasa Indonesia[8].

Penelitian berjudul "Penerapan Teks Mining dan *Cosine Similarity* untuk Menentukan Kesamaan Dokumen Skripsi" membahas upaya mencegah plagiarisme dalam pembuatan skripsi melalui digitalisasi dokumen. Penelitian ini menggunakan algoritma *cosine similarity* untuk mengukur kesamaan antara dokumen skripsi yang ada di perpustakaan Prodi Teknik Informatika Universitas Muhammadiyah Maluku Utara. Dengan membandingkan 30 judul skripsi lama dan 5 judul skripsi baru, penelitian ini menunjukkan bahwa nilai *cosine similarity* terbesar berkisar antara 0,1 hingga 0,49. Hasil ini menunjukkan bahwa metode *cosine similarity* efektif dalam mendeteksi kemiripan antara dokumen, membantu mencegah praktik plagiarisme dalam pembuatan skripsi[9].

Dari tinjauan pustaka yang telah dibahas, dapat dilihat bahwa terdapat beberapa area yang masih memiliki kekurangan atau gap dalam penelitian terkait. Untuk lebih memahami kekurangan-kekurangan ini, berikut disajikan tabel gap penelitian yang merangkum perbedaan dan kekurangan dari berbagai penelitian yang relevan.

Tabel 1.1 Gap Penelitian

No.	Judul Penelitian	Temuan Utama	Gap Penelitian
1.	Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM	LSTM efektif dalam mengklasifikasikan emosi dari <i>tweet</i> dengan akurasi 80,42%. Dominasi sentimen "bingung".	Fokus pada analisis sentimen, tanpa mengembangkan sistem rekomendasi atau menggunakan <i>cosine similarity</i> .
2.	Uji Kemiripan Kalimat Judul Tugas Akhir dengan Metode <i>Cosine Similarity</i> dan Pembobotan TF-IDF	<i>Cosine similarity</i> dan TF-IDF baik untuk mengevaluasi judul tugas akhir. 43% judul tidak layak.	Terbatas pada evaluasi kemiripan judul tugas akhir, tanpa LSTM atau pengembangan sistem rekomendasi.
3.	Prediksi Genre Film Dengan Klasifikasi Multi Kelas Sinopsis Menggunakan Jaringan LSTM	LSTM memprediksi genre film dengan akurasi 98% menggunakan NLP.	Hanya menggunakan LSTM untuk prediksi genre film, tanpa <i>cosine similarity</i> atau penerapan dalam rekomendasi layanan.
4.	Meningkatkan Akurasi <i>Long-Short</i>	<i>GloVe</i> meningkatkan akurasi LSTM dari	Fokus pada analisis sentimen tanpa

	<p><i>Term Memory</i> (LSTM) pada Analisis Sentimen Vaksin Covid-19 di Twitter dengan <i>GloVe</i></p>	<p>87% menjadi 89%. Menggunakan vektor kata untuk <i>embedding</i>.</p>	<p>pengembangan sistem rekomendasi atau penggunaan <i>cosine similarity</i>.</p>
5.	<p>Penerapan Algoritma <i>Cosine Similarity</i> dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta</p>	<p>TF-IDF dan <i>cosine similarity</i> efektif untuk FAQ PMB dengan akurasi 64.28%.</p>	<p>Hanya diterapkan untuk FAQ PMB, tanpa LSTM atau pengembangan sistem rekomendasi layanan berbasis <i>website</i>.</p>
6.	<p>Penerapan Metode LSTM dalam Pembuatan Sistem Pendeteksi Berita Palsu Berbahasa Indonesia</p>	<p>LSTM dan <i>Indobert</i> mendeteksi berita palsu dengan akurasi 94.2%.</p>	<p>Fokus pada deteksi berita palsu, tanpa <i>cosine similarity</i> atau pengembangan sistem rekomendasi berbasis layanan.</p>
7.	<p>Penerapan Teks Mining dan <i>Cosine Similarity</i> untuk Menentukan</p>	<p><i>Cosine similarity</i> efektif untuk mendeteksi kemiripan dokumen</p>	<p>Hanya menggunakan <i>cosine similarity</i>, tanpa LSTM atau pengembangan sistem</p>

	Kesamaan Dokumen Skripsi	skripsi. Nilai kemiripan antara 0,1 hingga 0,49.	rekomendasi berbasis <i>website</i> .
--	--------------------------	--	---------------------------------------

Dari tinjauan pustaka dan tabel gap penelitian, dapat disimpulkan bahwa meskipun metode LSTM dan *cosine similarity* efektif dalam aplikasi seperti analisis sentimen dan deteksi plagiarisme, masih ada kekurangan yang perlu diperbaiki. Penggunaan metode ini menunjukkan potensi besar untuk sistem rekomendasi, termasuk untuk *freelancer*, tetapi perlu penyesuaian dan teknik terbaru agar hasilnya lebih baik.

#### 1.4. Data Penelitian

Dalam penelitian ini, data yang digunakan untuk mengembangkan sistem rekomendasi *freelancer* adalah *dataset* yang telah dikumpulkan. Berikut penjelasan mengenai bentuk, jumlah dan sumber *dataset* yang digunakan.

##### 1) Bentuk Data

*Dataset* ini berisi paragraf kebutuhan pengguna yang diambil dari media sosial seperti Twitter dan platform sejenis. Selain itu, data ini juga mencakup kata kunci yang telah diekstraksi dari teks tersebut. Berikut penjelasannya:

- a) Paragraf Kebutuhan Pengguna: Data dalam format teks yang mencakup deskripsi kebutuhan atau permintaan pengguna untuk layanan *freelancer*.

b) Kata Kunci: Data dalam format teks yang mewakili kata kunci yang telah diekstraksi dari paragraf kebutuhan pengguna secara manual.

2) Jumlah Data

Jumlah *dataset* yang digunakan sekitar kurang lebih 2000 data.

3) Sumber Data

Media sosial seperti Twitter dan platform sejenis di mana pengguna mengungkapkan kebutuhannya untuk layanan *freelancer*. Data ini dikumpulkan melalui API atau teknik *crawling* untuk memperoleh teks kebutuhan pengguna serta kata kunci yang relevan.

Data ini akan digunakan untuk melatih model LSTM dalam mengekstraksi kata kunci dari teks kebutuhan pengguna dan untuk mencocokkan kebutuhan pengguna dengan *freelancer* yang sesuai menggunakan metode *cosine similarity*. Tujuannya adalah untuk mengembangkan sistem rekomendasi yang dapat memberikan rekomendasi lebih akurat sesuai kebutuhan pengguna.