



Evaluasi Estimasi Biaya Perangkat Lunak melalui Ekstraksi Katalog Dari Dokumen Spesifikasi Kebutuhan

Luqman Fanani¹, Andi Ikmal Rachman², Suriansyah B³, Gita Pratiwi⁴, Agus Halid⁵, Andi Alviadi Nur Risal⁶

Program Studi Sistem Informasi, Universitas Almarisah Madani^{1,2,3,4,5}

Program Studi Sistem Informasi, Universitas Megarezky⁶

Jl. Perintis Kemerdekaan Km 13, 7, Paccerakkang, Biringkanaya, Makassar, Sulawesi Selatan, Indonesia^{1,2,3,4,5}

Jl. Antang Raya No. 43, Manggala, Makassar, Sulawesi Selatan, Indonesia⁶

luqmanfmz@univeral.ac.id*¹, andiikmal@univeral.ac.id², suriansyahb@univeral.ac.id³,

Gitapратиwi@univeral.ac.id⁴, agushalid@univeral.ac.id⁵, andialviadinurrisal@unimerz.ac.id⁶.

Kata Kunci :

Effort
Estimation, SRS
Word2vec
Cosine similarity
Function Point

ABSTRAK

Estimasi biaya pembuatan perangkat lunak adalah tahap awal yang penting dalam siklus pengembangan perangkat lunak. Proses ini memerlukan analisis proyek yang teliti, dengan mempertimbangkan berbagai faktor - faktor yang mempengaruhi biaya dan waktu penyelesaian seperti kesalahan dalam melakukan identifikasi awal perangkat lunak seperti apa yang akan di bangun dan pemanfaatannya. Salah satu tantangan utama dalam penganggaran adalah kurangnya harga acuan yang jelas, yang sering kali mengakibatkan penggunaan data historis sebagai dasar perhitungan. Penelitian ini mengusulkan kombinasi metode untuk meningkatkan akurasi dan keandalan estimasi biaya, termasuk text summarization dan word2vec untuk analisis dan pembobotan kalimat, serta ekstraksi katalog untuk mengidentifikasi dokumen SRS sebagai fitur sistem, termasuk fitur ambiguitas. Tujuannya adalah menyediakan alat bantu yang lebih efektif untuk pembuatan penganggaran proyek perangkat lunak di masa mendatang, memastikan estimasi biaya yang sesuai dengan kompleksitas proyek dan penugasan tenaga ahli yang tepat. Dengan metode ini, diharapkan perusahaan dapat mengurangi risiko kesalahan perhitungan dan penugasan tenaga ahli yang tidak tepat, sehingga menghindari kerugian finansial dan penundaan proyek

Keywords

Effort
Estimation, SRS
Word2vec
Cosine similarity
Function Point

ABSTRACT

Software cost estimation is an important early stage in the software development cycle. This process requires careful analysis of the project, taking into account various factors that affect cost and time to completion such as errors in the initial identification of what kind of software will be built and its utilization. One of the main challenges in budgeting is the lack of clear reference prices, which often results in the use of historical data as the basis for calculations. This research proposes a combination of methods to improve the accuracy and reliability of cost estimation, including text summarization and word2vec for sentence analysis and weighting, and catalog extraction to identify SRS documents as system features, including ambiguity features. The goal is to provide a more effective tool for future software project budgeting, ensuring cost estimation that matches the complexity of the project and proper assignment of experts. With this method, it is expected that companies can reduce the risk of miscalculation and inappropriate assignment of experts, thereby avoiding financial losses and project delays.

---Jurnal JISTI @2024---



PENDAHULUAN

Dalam beberapa dekade terakhir, perkembangan teknologi telah signifikan mempengaruhi ukuran dan fungsionalitas perangkat lunak, menjadikan estimasi perangkat lunak sebagai elemen krusial dalam siklus hidup pengembangan perangkat lunak. Estimasi perangkat lunak merupakan tugas penting yang harus dilakukan sebelum tahap perancangan dimulai dan sering berlanjut sepanjang siklus perangkat lunak. Proses ini membantu dalam menghasilkan estimasi yang akurat terkait biaya dan hasil yang diharapkan, serta mengidentifikasi effort yang diperlukan untuk pengembangan aplikasi. Dengan estimasi yang tepat, alokasi sumber daya dapat dioptimalkan sehingga proyek dapat diselesaikan tepat waktu sesuai dengan anggaran yang ditetapkan (Mz et al., 2023).

Development life Cycle pengembangan perangkat lunak berbasis proyek sangat mempengaruhi oleh prediksi biaya pengembangan yang akurat. Terdapat berbagai model yang digunakan untuk estimasi biaya perangkat lunak. Dengan mendasarkan perhitungan pada spesifikasi kebutuhan perangkat lunak, fitur-fitur perangkat lunak dapat diekstraksi untuk mengestimasi biaya yang diperlukan untuk pengembangan. (Butt et al., 2022).

Penelitian terdahulu juga menerapkan ekstraksi fitur dokumen penawaran untuk mengatasi masalah efisiensi biaya, yang secara tradisional dihitung menggunakan metode perkiraan berdasarkan penilaian para ahli. Studi ini mengajukan ekstraksi fitur berbasis Function Point dan menguji metode tersebut pada 10 studi kasus dari berbagai industri. Hasil pengujian menunjukkan peningkatan efisiensi analisis biaya hingga 70%, yang tervalidasi secara empiris, membuktikan efektivitas pendekatan berbasis Function Point dibandingkan metode tradisional. Pendekatan ini memungkinkan estimasi biaya yang lebih terukur dan konsisten, mengurangi ketergantungan pada penilaian subjektif ahli, dan meningkatkan efisiensi operasional dalam proses penawaran dan perhitungan biaya (Delong Han et al., n.d.).

Penelitian ini bertujuan untuk menganalisis dan otomatisasi penganggaran yang ideal berdasarkan prinsip keilmuan Software engineering. Penulis menggunakan dokumen Software Requirement Specification (SRS) sebagai basis perhitungan, dikarenakan dokumen ini merupakan dasar perancangan sistem yang akan dikembangkan dan dapat digunakan sebagai tolak ukur perhitungan biaya perangkat lunak. Prosesnya melibatkan ekstraksi fitur dari setiap fungsionalitas dalam SRS, yang kemudian dihitung menggunakan metode Function Point. Metode ini diharapkan memberikan pendekatan yang lebih sistematis dalam menentukan besaran biaya dan alokasi tenaga ahli, dibandingkan metode metode perkiraan kasar (guesstimate).

KAJIAN PUSTAKA

1. Dokumen spesifikasi kebutuhan perangkat lunak

Dokumen spesifikasi kebutuhan perangkat lunak atau Software Requirements Specification (SRS) adalah dokumen yang dibuat dan berisi gambaran mendetail atas semua aspek dari perangkat lunak yang akan dibangun. The U.S. Institute of Electrical and Electronic Engineer (IEEE) adalah salah satu provider standardisasi terbaik saat ini yang juga telah mengembangkan panduan struktur dokumen dalam penulisan spesifikasi kebutuhan perangkat lunak. Meskipun 20 panduan ini lebih tepat untuk sistem militer dan sistem kontrol, karena memang dikembangkan Bersama dengan U.S. Department of Defense (DoD), namun panduan tersebut bisa dijadikan sebagai dasar untuk pengembangan penulisan SRS yang tentunya disesuaikan dengan kebutuhan perusahaan atau organisasi (Karl Wiegers, 2013).

2. Pemrosesan bahasa natural (NLP)

Bahasa natural masih menjadi pilihan utama bagi mayoritas pengembang perangkat lunak untuk mendefinisikan kebutuhan, meskipun sebenarnya telah tersedia bahasa formal beserta perangkat lunak



pendukungnya yang dianggap lebih presisi dalam representasi kebutuhan dan meminimalisasi ambiguitas yang mungkin terjadi. Oleh karena itu, pemrosesan bahasa natural (Natural Language Processing) dibutuhkan dalam setiap proses otomatisasi yang melibatkan dokumen teks seperti halnya dalam penelitian ini. Penggunaan NLP dalam suatu sistem bukan berarti kemampuan untuk memahami teks secara sepenuhnya, tetapi lebih bertujuan untuk melakukan ekstraksi konsep yang terdapat pada suatu dokumen (Shamshiri et al., 2024).

3. Word2vec

Word2vec adalah sebuah algoritma yang digunakan untuk membuat representasi vektor dari kata-kata dalam bentuk angka, yang juga disebut sebagai "word embeddings". Representasi vektor ini memungkinkan komputer untuk memahami makna kata dan hubungannya dengan kata-kata lain dalam sebuah kalimat atau dokumen. Setelah dilatih, representasi vektor kata yang dihasilkan oleh model word2vec dapat digunakan untuk berbagai tugas pemrosesan bahasa alami, seperti pemodelan bahasa, analisis sentimen, dan terjemahan mesin. Word2vec juga sangat berguna dalam memproses dataset yang memiliki banyak sekali kata-kata unik (Muhammad et al., 2021).

4. Algoritma cosine similarity

Cosine similarity berfungsi untuk membandingkan kemiripan antar dokumen, dalam hal ini yang dibandingkan adalah query dengan dokumen latih. Dalam menghitung cosine similarity, pertama yang dilakukan yaitu melakukan perkalian skalar antara query dengan dokumen kemudian dijumlahkan, setelah itu melakukan perkalian antara panjang dokumen dengan panjang query yang telah dikuadratkan, setelah itu di hitung akar pangkat dua (Rinjeni et al., 2024). Rumus dapat dilihat sebagai berikut :

Untuk dua vektor A dan B yang memiliki dimensi n:

$$A = (A_1, A_2, \dots, A_n) \quad (1)$$

$$B = (B_1, B_2, \dots, B) \quad (2)$$

Cosine similarity antara vektor A dan B didefinisikan sebagai:

$$\text{Cosine similarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} \quad (3)$$

Dengan demikian, rumus cosine similarity dapat ditulis secara lengkap sebagai:

$$\text{Cosine similarity}(A, B) = \frac{\sum_{i=1}^n A_i \cdot B_i}{\sqrt{\sum_{i=1}^n A_i^2} \cdot \sqrt{\sum_{i=1}^n B_i^2}} \quad (4)$$

5. Function points

Function Point Analysis (FPA) adalah metode untuk mengukur ukuran fungsional perangkat lunak berdasarkan fungsionalitas yang disediakan kepada pengguna. FPA mengukur perangkat lunak dalam unit yang disebut Function Points (FP) (Hu, 2024). Berikut adalah langkah-langkah dan rumus untuk menghitung Function Points:

a. Identifikasi komponen fungsional:

- External Inputs (EI): Jumlah input eksternal ke sistem.
- External Outputs (EO): Jumlah output eksternal dari sistem.



- External Inquiries (EQ): Jumlah permintaan (query) yang mengakses data dari dalam sistem.
 - Internal Logical Files (ILF): Jumlah file logis internal yang digunakan oleh sistem.
 - External Interface Files (EIF): Jumlah file logis eksternal yang digunakan oleh sistem.
- b. Tentukan nilai bobot untuk masing-masing komponen
 Bobot diberikan berdasarkan kompleksitas masing-masing komponen (sederhana, sedang, kompleks).

Table 1. complexity

Komponen Fungsional	Sederhana	Sedang	Kompleks
External Input (EI)	3	4	6
External Output (EO)	4	5	7
External Inquiry (EQ)	3	4	6
Internal Logical File (ILF)	7	10	15
External Interface File (EIF)	5	7	10

- c. Hitung function point tidak tertimbang (Unadjusted Function Points - UFP):

$$UFP = \sum(\text{Jumlah Komponen} * \text{Bobot}) \quad (5)$$

- d. Hitung Faktor Penyesuaian (Value Adjustment Factor - VAF):

Evaluasi 14 karakteristik sistem umum, masing-masing dengan nilai dari 0 sampai 5, untuk menentukan tingkat pengaruhnya. Jumlah nilai dari semua karakteristik memberikan Total Degree of Influence (TDI).

Faktor Penyesuaian (VAF):

$$VAF = 0.65 + (0.01 * TDI) \quad (6)$$

- e. Hitung Function Points (FP)

$$FP = UFP + VAF \quad (7)$$

METODE PENELITIAN

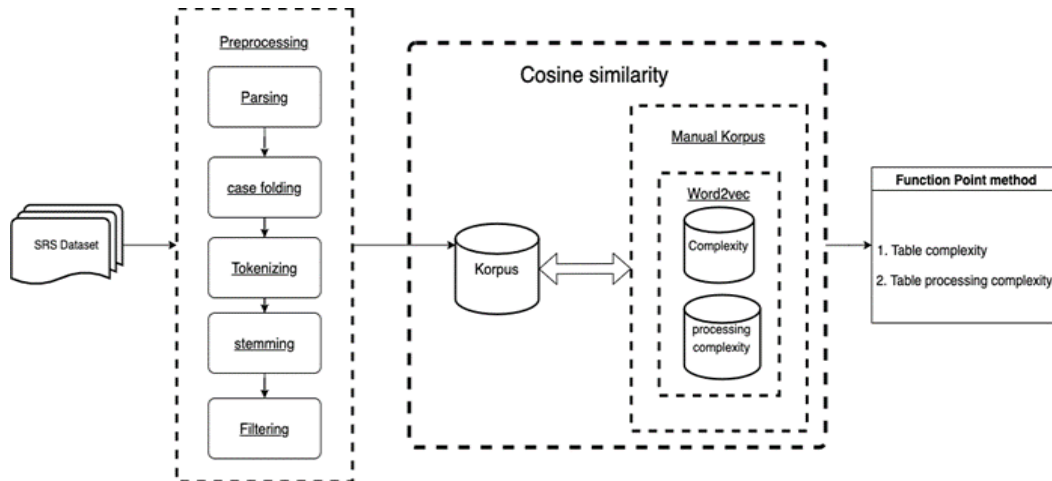
1. Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah kuantitatif yang digunakan untuk mendapatkan data yang terjadi saat ini. Dengan melakukan perancangan dan pengujian model sistem. Penelitian ini berfokus untuk mengestimasi nilai perangkat lunak secara elektronik (Manikavelan & Ponnusamy, 2020), pada sistem ini akan menghasilkan output satuan harga dan alokasi tenaga ahli yang ideal dalam satu siklus pengembangan suatu pengadaan perangkat lunak.



2. Perancangan Sistem

Pada rancangan sistem ini penulis akan melakukan penelusuran seluruh data primer dari seluruh seluruh dokumen resmi spesifikasi kebutuhan berdasarkan acuan IEEE penganggaran perangkat lunak yang telah terlaksanakan baik sumber dari pihak kontraktor atau pemerintah (Suhartoyo & Wijaya, n.d.). Adapun struktur blok diagram dari rancangan sistem ini seperti pada gambar dibawah ini :



Gambar 1. Blok Diagram

Dari dataset SRS yang diperoleh, dilakukan preprocessing menggunakan natural language processing diantaranya :

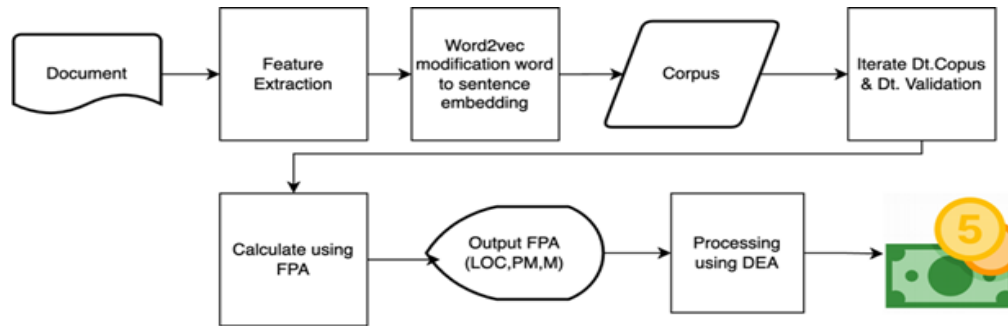
- a. Parsing
- b. Case folding
- c. Tokenizing / lexical analysis
- d. Stemming
- e. Filter

Setelah data korpus yang di ciptakan oleh proses di atas selesai maka akan dihitung nilai kedekatannya menggunakan algoritma cosine similarity dan dihadapkan dengan data korpus yang diciptakan oleh penulis secara manual berdasarkan sumber dari data staff analysis sistem sebagai acuan pengukuran. Data korpus yang diciptakan melalui preprocessing dan diciptakan oleh penulis nantinya akan di embedding scheme bobot menggunakan algoritma word2vec. Dan jika menilai seberapa penting istilah / kalimat tersebut dalam korpus dihitung dengan log jumlah dokumen dengan istilah di dalamnya. Korpus atau dataset nantinya akan dibagi menjadi beberapa bagian diantaranya:

Korpus yang tercipta dari preprocessing dokumen spesifikasi kebutuhan

- a. Korpus yang disediakan sebagai pembandingan terbagi atas 2 dataset diantaranya
- b. Korpus dikhususkan untuk tabel complexity akan diciptakan menjadi 3 bagian sesuai dengan aturan penggunaan metode function point yaitu low, medium, high
- c. korpus untuk tabel process complexity.

Setelah hasil pengukuran telah di dapatkan dari pemrosesan di atas maka metode function point akan menjalankan proses hingga mendapatkan nilai LOC (line of code) hingga menghasilkan estimasi berapa personel , berapa bulan pengerjaan hingga biaya yang harus disiapkan. Adapun detail alur rancangan sistem seperti gambar dibawah ini.



Gambar 2. Detail Alur System

HASIL DAN PEMBAHASAN

1. Implementasi Sistem

Implementasi sistem didasarkan pada hasil rancangan yang telah disusun pada tahap sebelumnya, dengan demikian dapat disimpulkan bahwa Langkah-langkah dalam menerapkan sistem telah sesuai dengan rancana awal mendapatkan nilai ideal dari satu penggaran perangkat lunak, sistem ini dibangun menggunakan IDE jupyter notebook yang dijalan secara parsial (per-blok).

Output Sistem estimasi harga perangkat lunak

```

Deteksi Fungsionalitas SRS
4 fitur sistem mengiputkan data asesment. - pemberitahuan asesment yang akan kadaluarsa. - pencarian berkas asesment. - import data kepegawaian.

Similaritas data uji dan corpus
4 fitur sistem mengiputkan data asesment. TERDEKAT DENGAN tersedia fitur pencarian data. SENILAI 0.5070925528371
1 DENGAN LABEL queries-normal DAN BOBOT 7
pemberitahuan asesment kadaluarsa. TERDEKAT DENGAN informasi barang rusak kadaluarsa. SENILAI 0.4472135954999579
DENGAN LABEL queries-normal DAN BOBOT 4
pencarian berkas asesment. TERDEKAT DENGAN pencarian berkas berdasarkn kunci. SENILAI 0.6780203932499369 DENGAN
LABEL queries-normal DAN BOBOT 7
import data kepegawaian. TERDEKAT DENGAN import ekspor data geojson. SENILAI 0.6780203932499369 DENGAN LABEL
berkas-normal DAN BOBOT 7

Output sistem
allocation resource : 3.33 Person
Working time : 4.479910446764304 Month
=====
DETAILS
Planning divisions take time 0.4995 Month requires personnel 0.6719865670146455 Person
Analysis Division takes time 0.666 Month requires personnel 0.8959820893528608 Person
Divisi desain membutuhkan waktu 1.1655 Month requires personnel 1.5679686563675064 Person
Divisi implementasi membutuhkan waktu 0.999 Month requires personnel 1.343973134029291 Person
=====
Planning Cost : Rp. 3225536
Analysis Cost : Rp. 4300714
System Design Costs : Rp. 7526250
Implementation Cost : Rp. 6451071
=====
Total base budget : Rp. 21503570
    
```

Gambar 3. Output Sistem

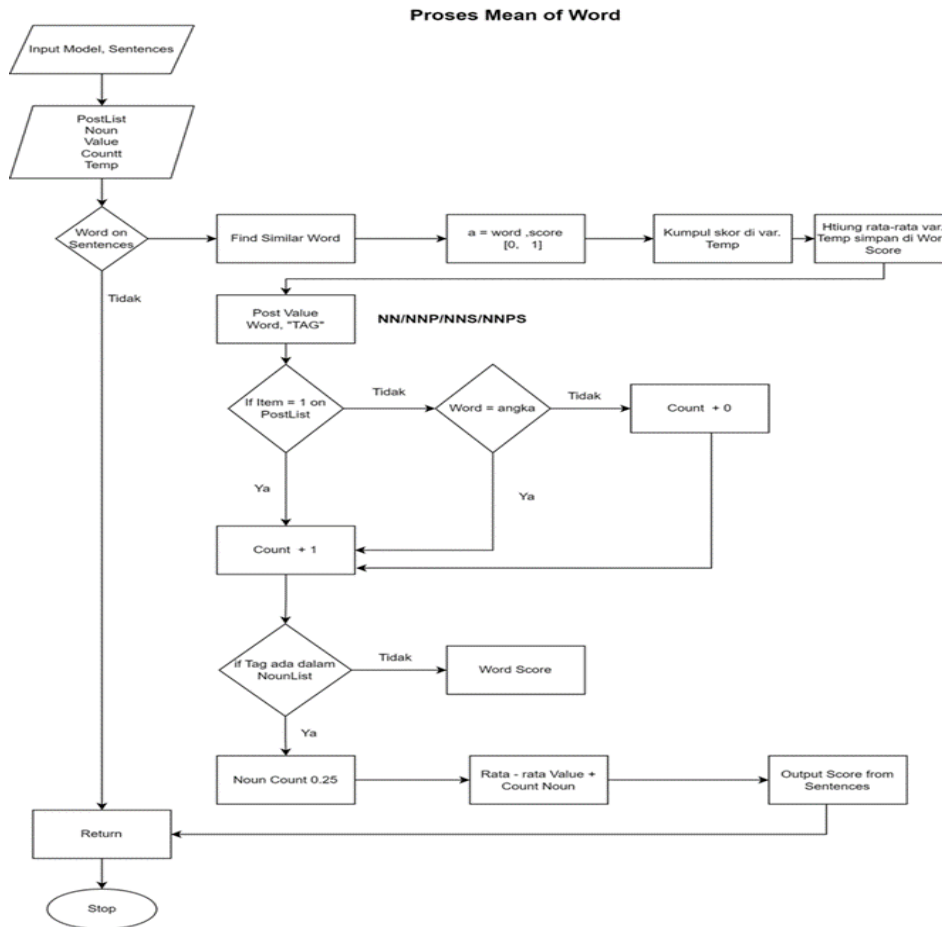
2. Preprocessing dokumen

Penelitian ini akuisisi data meliputi ekstraksi fitur. jumlah data yang diperoleh dan digunakan adalah 100 data penawaran perangkat lunak dan menghasilkan ekstraksi fitur fungsionalitas sebanyak 5000 data disimpan dalam format .csv. Atribut yang digunakan adalah bobot dari satu kalimat fungsionalitas beserta labling skoring manual dari hasil observasi pada 10 tenaga ahli dengan minimum pengalaman 10 tahun pada software development. Semua data didasarkan pada perusahaan di Kota Makassar yang di aplikasikan di beberapa daerah di wilayah indonesia.berikut adalah contoh parsial dataset yang digunakan sebagai training dibawah ini.



Table 2. Hasil ekstraksi fitur

Sentences	Fp-Labeling	Weight/Distance	Scoring FP
Pengecekan Dokumen Pemohon.	Queries-Normal	1.017.050.777.375.690	5
Mengelola Berkas Prose Staf Lapangan.	Berkas-Rendah	15.292.736.458.281.600	2
Pemohon Mengunggah Berkas Smartphone Android File Pembayaran Pajak Bumi Terbaru Alamat.	Masukan-Rendah	3.010.222.935.800.750	1
Petugas Lapangan Penandaan Pengukuran Lokasi Citra Satelit Digital.	Masukan-Rendah	2.265.777.332.418.490	1
Cari Berdasarkan Nik.	Queries-Rendah	10.185.830.052.942.000	2
Pelaporan Kunjungan Petugas Lapangan.	Queries-Normal	12.919.473.803.043.300	4
Pelaporan Data Telah Diproses .	Queries-Normal	12.707.143.800.954.000	4
Monitoring Aktivitas Sistem Dipantau Operator.	Keluaran-Normal	15.294.827.210.406.400	5
Pendaftaran Data Pemohon Izin Membangun Bangunan.	Masukan-Normal	17.638.528.040.477.200	4
Penugasan Pegawai Lapangan Rumah Pemohon.	Masukan-Rendah	15.287.037.380.039.600	1
Melaporkan Hasil Kegiatan Sistem.	Masukan-Normal	12.648.813.354.969.000	4
Monitoring Data Pemohon Pegawai Lapangan .	Keluaran-Normal	15.280.871.112.431.700	7
Pelaporan Data Telah Belum Proses.	Berkas-Rendah	1.260.736.513.733.860	1
Unduh Data Kunjungan.	Antarmuka-Rendah	10.098.481.640.219.600	1
Informasi Destinasi Wisata.	Keluaran-Rendah	10.430.967.096.239.300	2
Rekomendasi Wisata Berdasarkan Jenis Kunjungan.	Queries-Tinggi	15.128.086.147.209.000	9
Informasi Lokasi Penginapan Disesuaikan Berdasarkan Anggaran Wisatawan.	Queries-Rendah	20.199.070.131.406.100	3
Manajemen Perjalanan Dibuatkan Otomatis Sistem.	Keluaran-Tinggi	15.140.694.774.687.200	9
Sistem Terintegrasi Aplikasi Arcgis Dinas Pariwisata.	Keluaran-Normal	17.591.258.046.882.400	7
Monitoring Kunjungan Wisatawan Lokasi.	Keluaran-Rendah	1.280.795.176.923.270	3
Pelaporan Rekomendasi Wisatawan.	Antarmuka-Normal	10.298.984.937.369.800	7
Pengiriman Informasi Secara Massal Wisatawan.	Antarmuka-Rendah	15.184.522.283.573.900	3
Preview Lokasi Video Gambar.	Keluaran-Tinggi	12.764.575.892.686.800	8
Merekomendasikan Penginapan Terdekat Termurah Lokasi Tujuan.	Queries-Normal	17.596.108.036.381.800	7



Gambar 4. Flowchart Modifikasi Word To Sentaces Pada Algoritma Word2vec

Pada flowchart di atas Proses ini dimulai dengan memasukkan data yang telah diekstraksi fiturnya. Data ini kemudian disimpan sementara dalam suatu variabel temporer yang memuat kalimat yang terdiri dari benda dan kata kerja, membentuk satuan kalimat fungsional. Setiap kalimat kemudian dianalisis untuk mengecek similaritasnya menggunakan metode cosine similarity. Untuk setiap kalimat yang telah terlabel dan berbobot nilai, jika kalimat tersebut merupakan kalimat yang valid, maka akan diberikan bobot tambahan sebesar 1. Apabila kalimat tersebut mengandung identitas kata kerja, bobot tambahan sebesar 0,25 juga akan diberikan. Hasil akhir dari analisis ini akan disimpan ke dalam dataset atau korpus untuk keperluan lebih lanjut.

3. Langkah-langkah Terperinci:

- a. Inisiasi Proses
 - Proses dimulai dengan inisiasi sistem
- b. Input Data Ekstraksi Fitur
 - Data hasil ekstraksi fitur dimasukkan ke dalam sistem.
- c. Penyimpanan Sementara dalam Variabel Temporer
 - Data yang telah diinput disimpan sementara dalam variabel temporer yang memuat kalimat dengan struktur benda dan kata kerja.
- d. Iterasi Melalui Kalimat
 - Sistem melakukan iterasi melalui setiap kalimat yang ada dalam variabel temporer.

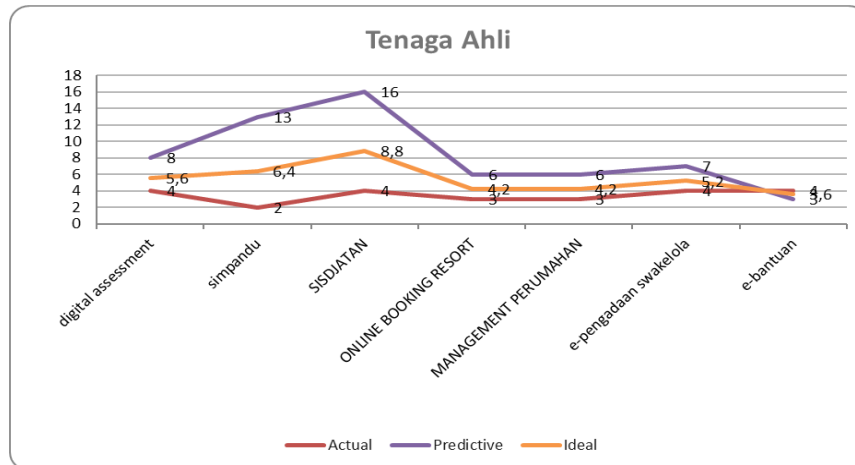


- e. Pengecekan Similaritas Menggunakan Cosine Similarity
 - Setiap kalimat dianalisis untuk mengecek similaritasnya menggunakan metode cosine similarity.
 - Kalimat Terlabel:
 - Jika kalimat terlabel, bobot tambahan sebesar 1 diberikan.
 - Kalimat Mengandung Kata Kerja:
 - Jika kalimat mengandung kata kerja, tambahan bobot sebesar 0,25 diberikan.
 - Kalimat Tidak Terlabel:
 - Jika kalimat tidak terlabel, iterasi dilanjutkan ke kalimat berikutnya tanpa penambahan bobot.
- f. Penyimpanan Hasil ke Dataset/Korpus
 - Hasil akhir dari analisis kalimat beserta bobotnya disimpan dalam dataset atau korpus.
- g. Terminasi Proses
 - Proses diakhiri setelah semua kalimat telah dianalisis dan disimpan

4. Analisis Hasil Pengujian Sistem

Untuk menguji dan mengevaluasi kinerja sistem yang dirancang, dilakukan analisis dengan membandingkan model perhitungan guesstimate dengan metode penelitian yang diusulkan oleh penulis, yaitu kombinasi penggunaan dokumen srs, algoritma Word2Vec, cosine similarity, Function Point (FP), COCOMO, dan metode distribusi effort. Pengujian menggunakan data penawaran pekerjaan di seluruh Indonesia, melalui tahapan pengumpulan, pra-pemrosesan, ekstraksi fitur, dan penghitungan estimasi. Hasil analisis diharapkan menunjukkan keakuratan dan efektivitas metode kombinasi dibandingkan pendekatan guesstimate tradisional, dengan evaluasi berdasarkan akurasi prediksi, efisiensi waktu, dan kemudahan penggunaan sistem.

Penelitian yang membandingkan estimasi harga perangkat lunak dengan metode guesstimate terhadap 80 proyek dari berbagai perusahaan pengembang di Kota Makassar menunjukkan kelemahan dalam aspek ilmiah, karena hanya mendasarkan perhitungannya pada harga perangkat lunak sebelumnya yang serupa dan alokasi tenaga ahli secara spekulatif. Sebaliknya, hasil pengujian menggunakan model yang diusulkan oleh penulis, yang menggabungkan berbagai metode dan algoritma seperti Word2Vec, cosine similarity, Function Point (FP), COCOMO, dan distributed effort berbasis dokumen spesifikasi kebutuhan (Software Requirement Specification, SRS), menunjukkan peningkatan akurasi dan penganggaran ideal. Model ini menggunakan analisis semantik untuk memahami konteks proyek, dan teknik estimasi terukur untuk mengalokasikan sumber daya secara optimal, menghasilkan estimasi yang lebih ilmiah dan ideal. Tujuh sampel data perbandingan antara kedua metode, yang divisualisasikan dalam grafik dan tabel, menunjukkan perbedaan signifikan dalam hasil estimasi, dengan metode kombinasi yang diusulkan lebih unggul dalam memberikan estimasi biaya perangkat lunak yang lebih akurat dan berbasis dokumen spesifikasi kebutuhan. seperti pada visualisasi grafik dan table dibawah ini :

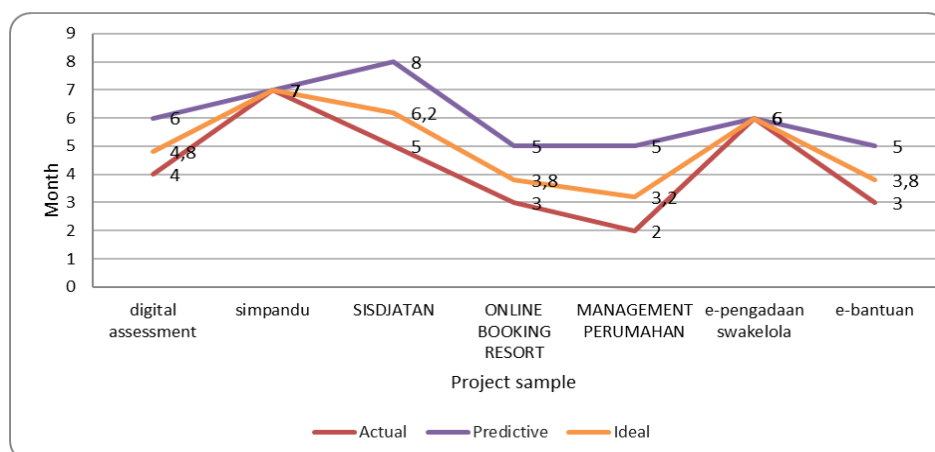


Gambar 5. Visualisasi penentuan alokasi tenaga ahli .

Pada Gambar 5, penulis memvisualisasikan hasil penerapan tujuh sampel data pekerjaan di sektor industri, yang menunjukkan perbedaan signifikan dalam distribusi dan pemanfaatan tenaga ahli, dengan waktu yang dibutuhkan hampir 20% lebih banyak dibandingkan metode guesstimate. Parameter seperti Actual of experts (Actual) dan Result from modification method on a research (Predictive) dari perbandingan tenaga ahli pada pekerjaan "simpandu" ini mengindikasikan bahwa tenaga ahli yang diperlukan untuk menyelesaikan perangkat lunak tersebut sangat minim, hanya sebanyak 2 orang.

Proses tersebut mengakibatkan overload tugas utama dan unit fungsi pada satu orang tenaga ahli, yang idealnya memerlukan 13 orang untuk memecah unit-unit pekerjaan menjadi lebih kecil agar kualitas perangkat lunak tetap terjaga, meskipun nilai ideal yang dihasilkan menggunakan metode statistik lebih tinggi, alokasi tenaga ahli sekitar lima hingga enam personil dengan garis regresi yang memiliki kemiringan positif ke arah aktual.

Oleh karena itu, hasil penelitian ini menegaskan pentingnya alokasi sumber daya yang lebih efektif dan seimbang untuk mencapai kualitas perangkat lunak yang optimal dalam lingkup pekerjaan "simpandu" di sektor industri.

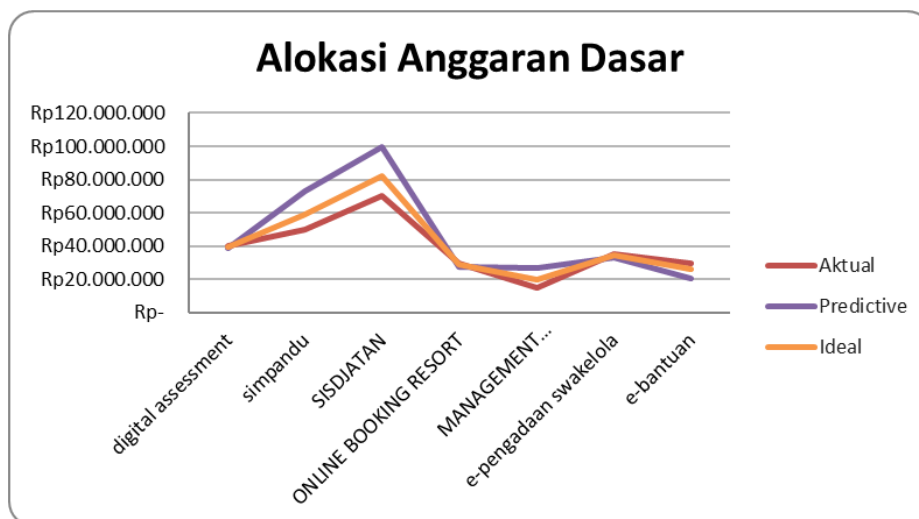


Gambar 6. Durasi visualisasi penentuan durasi pengerjaan (bulan).

Dalam Gambar 6, terdapat perbedaan yang cukup signifikan dalam distribusi pemanfaatan durasi waktu (bulan), khususnya pada pekerjaan "digital assessment" yang menunjukkan selisih empat bulan antara nilai aktual dan nilai prediktifnya, dengan nilai prediktif mencapai enam bulan. Melalui



pencarian nilai optimal menggunakan metode yang serupa, ditemukan bahwa durasi waktu optimal untuk pekerjaan tersebut adalah 4,8 bulan. Hasil ini menyoroti perbedaan yang substansial antara nilai aktual dan prediktif dalam hal durasi waktu pekerjaan, dimana pencarian nilai optimal memberikan rekomendasi durasi waktu yang lebih dekat dengan nilai aktual dibandingkan dengan nilai prediktif awal. Temuan ini memiliki implikasi penting dalam perencanaan dan alokasi sumber daya waktu untuk proyek "digital assessment," menegaskan kebutuhan akan metode evaluasi yang akurat dalam merencanakan dan mengeksekusi pekerjaan di sektor terkait.



Gambar 7. Grafik nilai aktual dengan metode yang diusulkan

Hasil perbandingan antara output yang diperoleh dari metode guesstimate dan metode function point yang telah dimodifikasi menunjukkan nilai dalam mata uang Rupiah yang mendekati nilai sebenarnya. Rata-rata persentase margin kesalahan terendah adalah sekitar 3%, sementara nilai tertinggi mencapai 46%. Tabel perbandingan hasil dari 7 sampel kegiatan memperlihatkan variabilitas dalam ketepatan estimasi antara kedua metode, dengan beberapa proyek menghasilkan perbedaan yang signifikan antara estimasi dan nilai sebenarnya.

Table 3. Komparasi persentase margin pada 2 metode beserta nilai aktual

Project	GE	FP	Margin (GE & FP)	Aktual
digital assessment	Rp. 40.000.000	Rp. 39.000.000	3%	Rp. 39.600.000,0
simpandu	Rp. 50.000.000	Rp. 73.000.000	-32%	Rp. 59.200.000,0
SISDJATAN	Rp. 70.000.000	Rp. 100.000.000	-30%	Rp. 82.000.000,0
ONLINE BOOKING RESORT	Rp. 30.000.000	Rp. 27.750.000	8%	Rp. 29.100.000,0
MANAGEMENT PERUMAHAN	Rp. 15.000.000	Rp. 27.000.000	-44%	Rp. 19.800.000,0
e-pengadaan swakelola	Rp. 35.000.000	Rp. 33.500.000	4%	Rp. 34.400.000,0
e-bantuan	Rp. 30.000.000	Rp. 20.500.000	46%	Rp. 26.200.000,0



Tabel 3 merupakan visualisasi perbandingan harga di mana beberapa sampel diharapkan mendapatkan nilai awal yang lebih besar, sementara yang lain seharusnya lebih kecil dari yang dihitung menggunakan metode guesstimate. Perbandingan ini diharapkan menjadi nilai acuan bagi pengembangan beberapa metode yang digunakan oleh penulis, dengan dampak yang signifikan terhadap investasi pekerjaan dan alokasi sumber daya tenaga ahli serta waktu yang dibutuhkan. Terdapat perbedaan sekitar \pm Rp. 23.000.000 antara harga aktual yang telah terimplementasi dan harga menurut metode function point yang dimodifikasi. Function point juga tidak selalu menghasilkan anggaran yang jauh lebih tinggi dari perkiraan, seperti pada pekerjaan "online booking resort" yang memiliki perbedaan nominal referensi minimal sebesar Rp. 3.000.000. Namun, terdapat perbedaan yang signifikan dalam perhitungan kebutuhan tenaga ahli dan waktu yang dibutuhkan dalam pengembangan perangkat lunak. Tabel 12 disediakan sebagai penyajian data yang lebih detail beserta informasi selisih anggaran, dimana nilai persentase positif menunjukkan perhitungan guesstimate yang terlalu besar yang seharusnya lebih efisien, sementara persentase negatif menandakan perhitungan guesstimate yang terlalu rendah dan berpotensi menimbulkan kerugian bagi pengembang perangkat lunak dibandingkan dengan metode function point yang konsep perhitungannya lebih detail dan ilmiah baik dalam perhitungan anggaran maupun alokasi sumber daya manusia.

SIMPULAN DAN SARAN

Sistem yang dikembangkan mampu menentukan besaran perangkat lunak serta alokasi tenaga ahli berdasarkan dokumen spesifikasi kebutuhan sebagai dasar perhitungan. Hasil estimasi menunjukkan bahwa metode perhitungan yang diusulkan dalam penelitian ini sangat relevan dan mengacu pada parameter keilmuan rekayasa perangkat lunak. Berdasarkan observasi lapangan pada 10 perusahaan dan pengumpulan data dari 40 proyek perangkat lunak, serta analisis dari 7 pengujian pekerjaan secara acak, ditemukan bahwa persentase kesalahan hitung berada pada margin di bawah 50 persen. Hal ini mengindikasikan bahwa sistem ini dapat memberikan estimasi yang akurat dan dapat diandalkan dalam mengukur ukuran perangkat lunak serta alokasi sumber daya manusia yang dibutuhkan untuk pengembangannya.

Untuk penelitian lebih lanjut, pengembangan sistem ini dapat diperluas dengan menentukan alokasi tenaga ahli secara lebih spesifik, seperti merinci spesifikasi tenaga ahli yang diperlukan dalam pengembangan perangkat lunak. Hal ini dapat mencakup identifikasi keahlian dan keterampilan yang diperlukan untuk setiap fase pengembangan, serta estimasi waktu yang diperlukan untuk setiap tugas yang dilakukan oleh tenaga ahli. Dengan demikian, sistem dapat memberikan estimasi yang lebih rinci dan akurat mengenai jumlah dan jenis tenaga ahli yang dibutuhkan dalam suatu proyek perangkat lunak, yang dapat menjadi panduan yang berharga bagi pengembang dalam merencanakan dan mengelola sumber daya manusia secara efektif.

DAFTAR PUSTAKA

- Butt, S. A., Khalid, A., Ercan, T., Ariza-Colpas, P. P., Melisa, A.-C., Piñeres-Espitia, G., De-La-Hoz-Franco, E., Melo, M. A. P., & Ortega, R. M. (2022). A software-based cost estimation technique in scrum using a developer's expertise. *Advances in Engineering Software*, 171, 103159. <https://doi.org/10.1016/j.advengsoft.2022.103159>
- Delong Han, Xungang Gu, Chengpeng Zhang, & Gang Li. (n.d.). *Research on Structured Extraction Method for Function Points Based on Event Extraction* _ Overview of research on function point



- Open Knowledge Maps.

Hu, Q. (2024). Change point analysis of functional variance function with stationary error. *Journal of Multivariate Analysis*, 202, 105311. <https://doi.org/10.1016/j.jmva.2024.105311>

Karl Wieggers, J. B. (2013). *Software_Requirements_3rd_Edition_compressed*.

Manikavelan, D., & Ponnusamy, R. (2020). Software quality analysis based on cost and error using fuzzy combined COCOMO model. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-020-01783-9>

Muhammad, P. F., Kusumaningrum, R., & Wibowo, A. (2021). Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews. *Procedia Computer Science*, 179, 728–735. <https://doi.org/10.1016/j.procs.2021.01.061>

Mz, L. F., Tahir, Z., & Suyuti, A. (2023). Development of Software Cost Estimation and Resource Allocation Using Natural Language Processing, Cosine Similarity and Function Point. 2023 International Conference on Digital Applications, Transformation & Economy (ICDATE), 1–6. <https://doi.org/10.1109/ICDATE58146.2023.10248788>

Rinjeni, T. P., Indriawan, A., & Rakhmawati, N. A. (2024). Matching Scientific Article Titles using Cosine Similarity and Jaccard Similarity Algorithm. *Procedia Computer Science*, 234, 553–560. <https://doi.org/10.1016/j.procs.2024.03.039>

Shamshiri, A., Ryu, K. R., & Park, J. Y. (2024). Text mining and natural language processing in construction. *Automation in Construction*, 158, 105200. <https://doi.org/10.1016/j.autcon.2023.105200>

Suhartoyo, hengki, & Wijaya, T. A. (n.d.). Rancangan Estimasi Biaya dengan Teknik COCOMO II dan Neuro Fuzzy Studi Kasus_ Sistem Informasi Rumah Sakit.