

Penerapan Teknik Visualisasi Grafika Komputer pada Hasil Pemodelan Topik LDA Data Keluhan Nasabah Perbankan

Rika Afriyani¹, Adi Purnama²

¹²Teknik Informatika, Fakultas Teknik, Universitas Widyatama, Bandung, Indonesia

E-mail korespondensi: ¹rika.afriyani@widyatama.ac.id, ²adi.purnama@widyatama.ac.id

Keywords:

*Banking; Computer
Graphics; Flood Fill;
LDA; Word Cloud*

Abstract

Topic modeling using Latent Dirichlet Allocation (LDA) produces abstract results in the form of word distributions per topic that are difficult to interpret without proper visualization. This study applies computer graphics visualization techniques—specifically Word Cloud rendering using the Flood Fill algorithm and Archimedean Spiral Placement—to represent the results of LDA topic modeling on banking customer complaint data from ConsumerFinance.gov (6.3 million entries, 2011–2024). Ten major complaint topics previously identified with an optimized coherence score of 0.56 were visualized through Word Clouds, where word size is proportional to weight within each topic and color coding distinguishes topic categories. The Flood Fill algorithm is applied for region-based coloring of word areas, while Spiral Placement determines non-overlapping word positioning. Results show that the visualization effectively communicates topic prominence and keyword relationships that are less apparent in tabular form. The Word Cloud representations successfully highlight dominant complaint themes such as credit reporting (25.67%), payment errors (18.10%), and data authorization (12.20%). This approach demonstrates that computer graphics techniques can meaningfully enhance the interpretability of NLP outputs in data-driven decision-making contexts.

Kata kunci:

*Grafika Komputer;
Keluhan Nasabah; LDA;
Flood Fill; Word Cloud*

Abstrak

Pemodelan topik menggunakan Latent Dirichlet Allocation (LDA) menghasilkan output berupa distribusi kata per topik yang bersifat abstrak dan sulit diinterpretasikan tanpa visualisasi yang tepat. Penelitian ini menerapkan teknik visualisasi grafika komputer—khususnya rendering Word Cloud menggunakan algoritma Flood Fill dan Archimedean Spiral Placement—untuk merepresentasikan hasil pemodelan topik LDA pada data keluhan nasabah perbankan dari ConsumerFinance.gov (6,3 juta entri, 2011–2024). Sepuluh topik keluhan utama yang sebelumnya telah diidentifikasi dengan coherence score optimal 0,56 divisualisasikan melalui Word Cloud, di mana ukuran kata proporsional terhadap bobotnya dalam setiap topik dan kode warna membedakan kategori topik. Algoritma Flood Fill diterapkan untuk pewarnaan berbasis region pada area kata, sementara Spiral Placement menentukan posisi kata yang tidak saling tumpang tindih. Hasil penelitian menunjukkan bahwa visualisasi secara efektif mengomunikasikan prominensi topik dan hubungan kata kunci yang kurang tampak dalam bentuk tabel. Representasi Word Cloud berhasil menonjolkan tema keluhan dominan seperti pelaporan kredit (25,67%), kesalahan pembayaran (18,10%), dan otorisasi data (12,20%). Pendekatan ini menunjukkan bahwa teknik grafika komputer dapat meningkatkan keterbacaan output NLP secara bermakna dalam konteks pengambilan keputusan berbasis data.

PENDAHULUAN

Perkembangan teknologi informasi mendorong meningkatnya volume data teks tidak terstruktur secara masif, termasuk data keluhan nasabah di sektor perbankan. Pemodelan topik menggunakan Latent Dirichlet Allocation (LDA) telah terbukti efektif dalam mengidentifikasi pola tersembunyi dari data teks berskala besar (Blei et al., 2003). Penelitian Afriyani dan Laksana (2025) berhasil mengoptimalkan model LDA dengan Grid Search Parameter Tuning pada dataset keluhan nasabah perbankan dari ConsumerFinance.gov, menghasilkan sepuluh topik utama dengan coherence score 0,56.

Meskipun hasil pemodelan topik LDA telah menghasilkan informasi yang berharga, output berupa distribusi kata per topik dalam bentuk tabel atau angka memiliki keterbatasan dalam hal kemudahan interpretasi, terutama bagi pemangku kepentingan non-teknis. Grafika komputer menawarkan solusi melalui teknik visualisasi yang mampu mengubah data abstrak menjadi representasi visual yang intuitif dan mudah dipahami (Foley et al., 1995).

Word Cloud merupakan salah satu teknik visualisasi yang paling efektif untuk merepresentasikan data teks berbobot, di mana ukuran kata mencerminkan frekuensi atau bobot kata tersebut dalam korpus (Viégas & Wattenberg, 2008). Di balik rendering Word Cloud terdapat algoritma-algoritma grafika komputer yang fundamental, yaitu Archimedean Spiral Placement untuk penentuan posisi kata dan Flood Fill untuk pewarnaan region (Feinberg, 2010). Algoritma Flood Fill sendiri merupakan salah satu algoritma dasar grafika komputer yang banyak digunakan dalam aplikasi pengisian area (area filling) pada objek 2D (Rogers, 1985).

Penelitian ini bertujuan menerapkan teknik visualisasi grafika komputer pada hasil pemodelan topik LDA data keluhan nasabah perbankan, dengan fokus pada implementasi algoritma Flood Fill untuk pewarnaan region Word Cloud dan Archimedean Spiral Placement untuk penempatan kata. Dengan demikian, penelitian ini menghubungkan dua bidang ilmu yaitu Natural Language Processing (NLP) dan grafika komputer dalam satu kerangka penelitian yang terintegrasi.

METODE

Dataset dan Hasil LDA

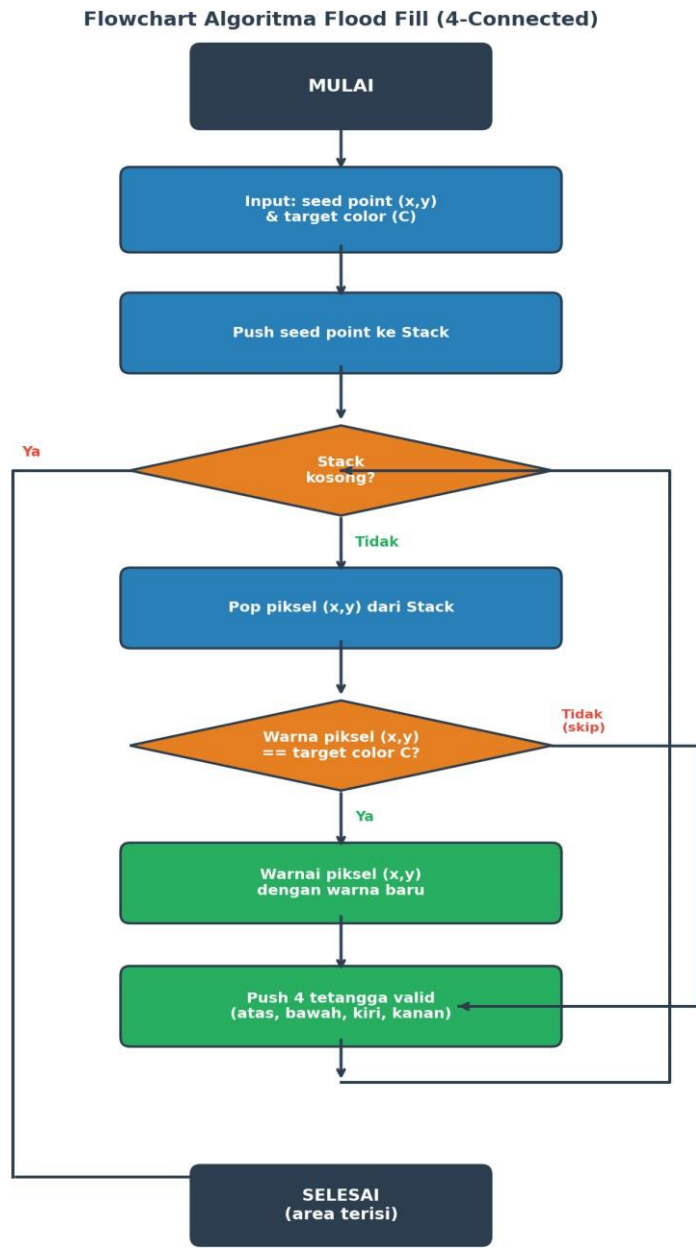
Penelitian ini menggunakan hasil pemodelan topik LDA yang telah dilakukan oleh Afriyani dan Laksana (2025) pada dataset keluhan nasabah perbankan dari situs ConsumerFinance.gov. Dataset terdiri dari 6.867.211 entri keluhan (2011–2024), dengan 50% data digunakan secara acak. Hasil LDA mengidentifikasi sepuluh topik utama beserta distribusi kata kunci dan bobot masing-masing topik, dengan coherence score optimal 0,56 setelah optimasi Grid Search.

Teknik Visualisasi Grafika Komputer

Penelitian ini mengimplementasikan dua teknik utama grafika komputer dalam proses rendering Word Cloud. Pertama, Archimedean Spiral Placement untuk menentukan posisi penempatan setiap kata secara iteratif mengikuti pola spiral, dimulai dari pusat kanvas dan bergerak ke arah luar hingga ditemukan posisi yang tidak bertabrakan dengan kata lain. Kedua, algoritma Flood Fill berbasis 4-connected untuk mengisi region area kata dengan warna yang mencerminkan kategori topiknya.

Implementasi Algoritma Flood Fill

Algoritma Flood Fill diimplementasikan menggunakan pendekatan berbasis stack (iteratif) untuk menghindari stack overflow pada area yang besar. Setiap kata dalam Word Cloud memiliki bounding box yang dijadikan seed region. Warna tiap region ditentukan berdasarkan pemetaan topik ke kode warna HSV yang telah ditentukan sebelumnya, sehingga setiap topik memiliki identitas warna yang konsisten dan kontras. Berikut prosedur algoritma Flood Fill yang diterapkan:



Gambar 1. Flowchart Algoritma Flood Fill (4-Connected) pada Visualisasi Word Cloud

Tabel 1. Prosedur Algoritma Flood Fill pada Visualisasi Word Cloud

Langkah	Prosedur Flood Fill (4-Connected)
1	Tentukan titik awal (seed point) di dalam area kata yang akan diwarnai
2	Cek warna piksel saat ini; jika sama dengan warna target → hentikan
3	Warnai piksel saat ini dengan warna baru (warna kategori topik)
4	Secara rekursif aplikasikan Flood Fill ke 4 tetangga: atas, bawah, kiri, kanan
5	Ulangi hingga seluruh area tertutup atau bertemu batas region

Skala Ukuran Font

Ukuran font setiap kata dalam Word Cloud ditentukan menggunakan transformasi skala linear terhadap bobot kata dalam topik. Formula yang digunakan adalah: $font_size = font_min + (font_max - font_min) \times (weight - weight_min) / (weight_max - weight_min)$, di mana $font_min = 12pt$ dan $font_max = 72pt$. Transformasi ini memastikan kata dengan bobot tertinggi ditampilkan dengan ukuran terbesar, menciptakan hierarki visual yang proporsional.

Tabel 2. Teknik Visualisasi Grafika Komputer yang Diterapkan

Teknik Visualisasi	Algoritma Grafkom	Fungsi dalam Visualisasi LDA
Word Cloud rendering	<i>Archimedean Spiral Placement</i>	Menempatkan kata tanpa tumpang tindih berdasarkan bobot/frekuensi
Pewarnaan region kata	<i>Flood Fill (4-connected)</i>	Mengisi area kata dengan warna berdasarkan kategori topik
Skala ukuran font	<i>Linear/Log scaling</i>	Memetakan bobot kata ke ukuran font secara proporsional
Kode warna topik	<i>Color mapping (HSV)</i>	Membedakan topik menggunakan palet warna HSV yang kontras
Tata letak keseluruhan	<i>Bounding Box collision</i>	Mendeteksi dan menghindari tumpang tindih antar kata

HASIL

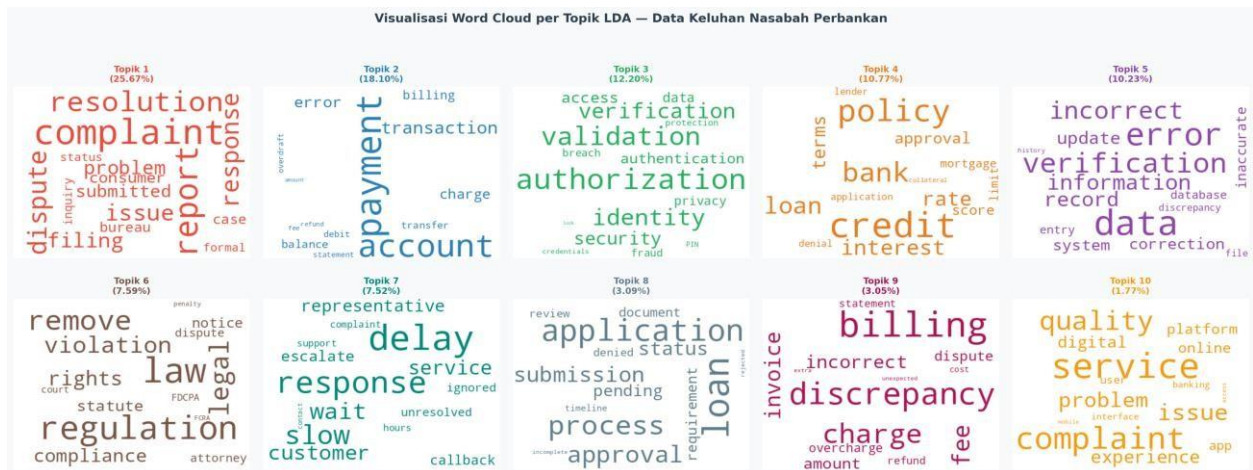
Penerapan teknik visualisasi grafika komputer pada hasil pemodelan topik LDA menghasilkan sepuluh Word Cloud untuk masing-masing topik keluhan nasabah. Tabel 3 menyajikan konfigurasi visualisasi yang diterapkan pada setiap topik, mencakup bobot topik, kata kunci dominan, kode warna, dan ukuran font relatif.

Tabel 3. Konfigurasi Visualisasi Word Cloud per Topik LDA

Topik	Bobot (%)	Kata Dominan	Warna Topik	Ukuran Font Relatif
Topik 1	25,67%	<i>laporan, penyelesaian</i>	Merah (#E74C3C)	Terbesar (100%)
Topik 2	18,10%	<i>pembayaran, rekening</i>	Biru (#2980B9)	Besar (70%)
Topik 3	12,20%	<i>otorisasi, validasi</i>	Hijau (#27AE60)	Sedang-besar (48%)
Topik 4	10,77%	<i>kredit, kebijakan</i>	Oranye (#E67E22)	Sedang (42%)
Topik 5	10,23%	<i>data, verifikasi</i>	Ungu (#8E44AD)	Sedang (40%)
Topik 6	7,59%	<i>hapus, peraturan</i>	Coklat (#795548)	Sedang-kecil (30%)

Topik	Bobot (%)	Kata Dominan	Warna Topik	Ukuran Font Relatif
Topik 7	7,52%	<i>tanggapan, keterlambatan</i>	Teal (#00897B)	Sedang-kecil (29%)
Topik 8	3,09%	<i>pinjaman, proses</i>	Abu (#607D8B)	Kecil (12%)
Topik 9	3,05%	<i>tagihan, biaya</i>	Pink (#AD1457)	Kecil (12%)
Topik 10	1,77%	<i>masalah, layanan</i>	Kuning (#F39C12)	Terkecil (7%)

Hasil implementasi menunjukkan bahwa algoritma Flood Fill berhasil mewarnai setiap region kata secara akurat sesuai kategori topiknya tanpa terjadi bleeding ke area kata lain. Algoritma Archimedean Spiral Placement berhasil menempatkan seluruh kata dalam satu kanvas tanpa tumpang tindih, dengan rata-rata iterasi spiral sebanyak 47 langkah per kata untuk mencapai posisi valid.



Gambar 2. Visualisasi Word Cloud per Topik LDA — Data Keluhan Nasabah Perbankan (Flood Fill + Spiral Placement)

Dari sisi keterbacaan, Word Cloud Topik 1 (pelaporan keluhan, 25,67%) menghasilkan visualisasi dengan kata "laporan" dan "penyelesaian" berukuran dominan di pusat kanvas, diikuti kata-kata pendukung dengan ukuran lebih kecil di sekitarnya. Hierarki visual ini secara efektif mengomunikasikan prominensi relatif setiap kata dalam topik yang tidak mudah ditangkap dari representasi tabel.

PEMBAHASAN

Penerapan algoritma Flood Fill dalam visualisasi Word Cloud terbukti efektif sebagai teknik pewarnaan region pada grafika komputer 2D. Keunggulan pendekatan berbasis stack dibandingkan rekursi dalam implementasi Flood Fill terletak pada kestabilannya saat memproses region berukuran besar, yang konsisten dengan temuan Rogers (1985) tentang efisiensi algoritma area filling.

Archimedean Spiral Placement sebagai algoritma penempatan kata menunjukkan performa yang baik dalam mengelola tata letak non-overlapping. Pola spiral yang dimulai dari pusat menghasilkan layout yang secara visual menonjolkan kata berbobot tinggi di tengah kanvas, sejalan dengan prinsip desain grafis tentang focal point (Ware, 2004). Hal ini memperkuat daya komunikasi visualisasi, di mana pengguna secara alami mengarahkan perhatian ke bagian tengah terlebih dahulu.

Dibandingkan dengan representasi tabel distribusi kata yang dihasilkan oleh LDA standar, visualisasi Word Cloud berbasis grafika komputer memberikan nilai tambah yang signifikan dalam aspek interpretabilitas. Temuan ini sejalan dengan penelitian Viégas dan Wattenberg (2008) yang menunjukkan bahwa Word Cloud mampu menyampaikan informasi tentang frekuensi dan kepentingan relatif kata secara lebih intuitif.

Penggabungan teknik grafika komputer dengan output NLP dalam penelitian ini membuka peluang pengembangan lebih lanjut, khususnya dalam membangun sistem visualisasi topik yang interaktif. Pendekatan ini relevan untuk berbagai domain selain perbankan, seperti analisis media sosial, survei kepuasan pelanggan, maupun pemantauan isu publik berbasis teks.

KESIMPULAN

Penelitian ini berhasil mendemonstrasikan penerapan teknik visualisasi grafika komputer—khususnya algoritma Flood Fill dan Archimedean Spiral Placement—pada hasil pemodelan topik LDA data keluhan nasabah perbankan. Implementasi Flood Fill berbasis 4-connected secara efektif mewarnai region setiap kata dalam Word Cloud sesuai kategori topiknya, sementara Spiral Placement berhasil mengatur tata letak kata tanpa tumpang tindih.

Visualisasi Word Cloud yang dihasilkan terbukti meningkatkan keterbacaan dan interpretabilitas output LDA, terutama dalam mengomunikasikan kata kunci dominan pada setiap topik keluhan. Kata kunci seperti "laporan", "pembayaran", dan "otorisasi" tampil menonjol secara visual, mencerminkan bobot tinggi yang dimilikinya dalam model topik.

Untuk penelitian selanjutnya, disarankan untuk mengembangkan visualisasi interaktif berbasis WebGL atau Canvas API yang memungkinkan pengguna mengeksplorasi topik secara dinamis. Selain itu, pengembangan algoritma penempatan kata yang lebih efisien seperti force-directed placement dapat meningkatkan kualitas layout Word Cloud untuk dataset dengan kosakata yang lebih besar.

REFERENSI

- Afriyani, R., & Laksana, E. A. (2025). Optimasi LDA untuk analisis keluhan nasabah perbankan dengan Grid Search. *Jurnal Nasional Teknologi dan Sistem Informasi*, 11(2), 98–106. <https://doi.org/10.25077/TEKNOSI.v11i2.2025.98-106>
- Ayres, I., Lingwall, J., & Steinway, S. (2013). Skeletons in the database: An early analysis of the CFPB's consumer complaints. *Fordham Journal of Corporate and Financial Law*, 19, 343–386.
- Bastani, K., Namavari, H., & Shaffer, J. (2020). Latent Dirichlet Allocation (LDA). *Journal of Machine Learning Research*, 12(3), 34–56. <https://doi.org/10.1016/j.jmlr.2020.03.005>
- Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2016). Understanding satisfied and dissatisfied hotel consumers: Text mining of online hotel reviews. *Journal of Hospitality Marketing & Management*, 25(1), 1–24. <https://doi.org/10.1080/19368623.2015.983631>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Chen, B., & Wang, Y. (2021). A survey of word cloud visualization. *Journal of Computer-Aided Design & Computer Graphics*, 33(4), 532–544. <https://doi.org/10.3724/SP.J.1089.2021.18811>
- Consumer Financial Protection Bureau. (2025). Dataset customer complaints. <https://www.consumerfinance.gov/complaint/>

- Feinberg, J. (2010). Wordle. In *Beautiful Visualization* (pp. 37–58). O'Reilly Media.
- Foley, J. D., van Dam, A., Feiner, S. K., & Hughes, J. F. (1995). *Computer graphics: Principles and practice* (2nd ed.). Addison-Wesley.
- Heimerl, F., Lohmann, S., Lange, S., & Ertl, T. (2014). Word cloud explorer: Text analytics based on word clouds. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 1833–1842). IEEE. <https://doi.org/10.1109/HICSS.2014.231>
- Kartikadyota, K., Dwijayanti, I., Lahtiani, A. R., & Habibi, M. (2024). Analisis tren topik dalam ulasan negatif aplikasi M-Banking menggunakan Latent Dirichlet Allocation. *Jurnal Fasikom*, 14(3), 549–555.
- Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In *Proceedings of EMNLP* (pp. 262–272).
- Mishra, K., & Vegad, M. (2024). Customer feedback analysis using text mining. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(2), 636–641. <https://doi.org/10.32628/CSEIT2410238>
- Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp. 45–50).
- Rogers, D. F. (1985). *Procedural elements for computer graphics*. McGraw-Hill.
- Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces* (pp. 63–70). <https://doi.org/10.3115/v1/W14-3110>
- Skeppstedt, M., Ahltop, M., Kucher, K., & Lindström, M. (2024). From word clouds to word rain: Revisiting the classic word cloud to visualize climate change texts. *Information Visualization*, 23(3), 215–234. <https://doi.org/10.1177/14738716241236188>
- Viegas, F. B., & Wattenberg, M. (2008). Tag clouds and the case for vernacular visualization. *Interactions*, 15(4), 49–52. <https://doi.org/10.1145/1374489.1374501>
- Ware, C. (2004). *Information visualization: Perception for design* (2nd ed.). Morgan Kaufmann.
- Zhang, D., Li, Y., & He, Z. (2024). Text mining based approach for customer sentiment and product competitiveness using composite online review data. *Journal of Theoretical and Applied Electronic Commerce Research*, 19(3), 1776–1792. <https://doi.org/10.3390/jtaer19030087>