JURNAL NUANSA INFORMATIKA
Volume 13 Nomor 1, Januari 2019
p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom

Pendeteksi Kesesuaian Format Laporan Skripsi dengan Panduan Penulisan menggunakan Machine Learning

Eko Travada Suprapto Putro
Fakultas Ilmu Komputer Universitas Nasional PASIM
Jl. Dakota No.8A Sukaraja Cicendo Bandung 40175 Telepon (022) 6072803
travada@pasim.ac.id / ekotravada@gmail.com

Abstrak

Machine Learning sebagai metode dalam system kecerdasan buatan yang mampu memodelkan data yang dimasukkan untuk kebutuhan di masa datang. Banyak aplikasi diterapkan seperti mengklasifikasi data, memprediksi hubungan antar data, merangking data, membaca pola data, membuat thriller film dan banyak implementasi lainnya. Dalam paper ini akan dibahas penggunaan keluaran struktur dari machine Learning digunakan untuk mendeteksi format skripsi mahasiswa apakah sudah sesuai dengan panduan skripsi. Berdasarkan hasil penelitian menunjukkan pemodelan struktur dari machine learning dapat digunakan untuk mendeteksi kesesuaian format baik pemodelan secara terpisah ataupun pemodelan secara gabungan.

Kata Kunci : Struktur Text, Machine Learning, Dokumen Skripsi

Abstract

Machine Learning is a method in artificial intelligence system that is able to model the data entered for future needs. Many applications are applied such as classifying data, predicting relationships between data, ranking data, reading data patterns, making movie thrillers and many other implementations. In this paper, we discuss the use of machine learning analysis result to detect the student's final project format whether it is in accordance with the final project guidelines or not. Based on research, it shows that the structural modeling of machine learning can be used to detect the suitability of the format both modeling separately and modeling in combination.

Key Word : Text Structure, Machine Learning, Final Project Documentation

1. Pendahuluan

Skripsi adalah mata kuliah terakhir yang ditempuh mahasiswa sebelum dinyatakan lulus dan menyandang program sarjana. Dalam penyusunan skripsi mahasiswa selain dibimbing oleh dosen juga mengikuti panduan yang telah ditentukan oleh masing masing masing fakultas atau program studi. Dalam panduan skripsi berisi aturan mengenai cara penulisan seperti spasi, jarak margin, jenis huruf dan sub bagian yang harus ada diantaranya latar belakang, maksud, tujuan identifikasi masalah. Kepatuhan mahasiswa mengikuti panduan yang telah ditetapkan menjadi salah satu komponen penilaian saat sidang skripsi.

Peranan dosen pembimbing adalah membimbing mahasiswa agar dapat menyusun laporan skripsi dengan baik. Format laporan skripsi tidak menjadi kewajiban dosen untuk memeriksanya karena setiap mahasiswa sudah

p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom

mendapat panduan yang dapat diunduh. Dengan demikian kepatuhan mengikuti format dikembalikan ke mahasiswa yang sedang menyusun skripsi. Dalam perjalanannya kepatuhan mahasiswa untuk mengikuti format yang sesuai panduan relatif rendah, hal ini dapat dilihat dari jajak pendapat Kepada dosen pembimbing dan dosen penguji skripsi. Dari hasil jajak pendapat sekitar 40 % mengatakan salah satu penghambat perbaikan dari laporan skripsi adalah lemahnya laporan dalam mengikuti panduan yang telah ditetapkan terutama di bab awal atau di bab I.

Bab I merupakan kunci agar penguji atau pembaca dapat memahami mengenai penelitian yang dilakukan. Banyak terjadi isi bab I beragam format dan tidak sesuai dengan panduan sehingga menurunkan nilaui mahasiswa saat sidang skripsi dan juga memakan waktu saat proses bimbingan untuk memperbaiki format laporan. Dampaknya waktu bimbingan untuk fokus ke isi laporan jadi berkurang. Dengan kondisi ini dibutuhkan suatu teknik yang mampu meneteksi perbedaan tersebut dengan cepat tanpa perlu melakukan pengecekan dokumen dengan dibaca per sub babnya satu persatu.

Salah satu tools yang bisa digunakan untuk menjaga format dari suatu laporan menggunakan latex. Latex dapat mendisain *template* dari format suatu laporan. Yang menjadi persoalan adalah bila panduan yang digunakan berubah maka diperlukan usaha yang cukup besar untuk merubah *template* tersebut.

Selain permasalahan yang telah disampaikan di paragraph diatas panduan pembuatan skripsi laporan umumnya secara periodik dalam jangka waktu tertentu diubah menyesuaikan dengan kebutuhan, beberapa panduan laporan skripsi kadang kala tidak terlalu kaku dalam menerapkan aturannya. Tapi walaupun demikian yang menjadi inti sub bagian dari suatu laporan tetap harus diikuti. Dengan kedinamisan dari suatu format laporan ini tentunya dibutuhkan suatu metode yang mampu mengakomodir perubahan pola tersebut.

Salah satu teknologi yang mampu beradaptasi terhadap suatu kondisi yang sering berubah adalah system cerdas. Riset mengenai system cerdas saat ini sudah diterapkan di hampir berbagai bidang dan salah satu teknologi di bidang system cerdas adalah machine Learning (mesin pembelajar). Mesin learning adalah suatu mesin yang mampu memodelkan data data yang dimasukkan untuk kemudian bisa membentuk model dari data tersebut (Wuillemin and Torti, 2012). Dengan kemampuan machine learning memodelkan data sesuai masukkan maka pada penelitian ini akan diuji cobakan penggunaan machine learning untuk mendeteksi kesesuaian laporan skripsi mahasiswa dengan panduan yang telah ditentukan.

Untuk penelitian awal di paper ini pengukuran kesesuaian dibatasi di bab I dan pola yang dilihat di kesesuaian sub Bab I. Untuk pengecekan margin dan spasi dapat menjadi penelitian berikutnya.

2. Teori Dasar

2.1.Struktur Text

p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom

Struktur Text adalah struktur yang bisa terbentuk dari text yang dimasukkan (Sebastiani, 2002). Dalam mendeteksi laporan skripsi maka dilakukan strukturisasi dari suatu skripsi untuk dapat dibandingkan dengan panduan apakah strukturnya sama atau berbeda. Dalam menstrukturisasi dari suatu dokumen secara formal bila Dokumen disimbolkan dengan D kemudian sejumlah dokumen disimbolkan dengan i maka ditulis D_i yang secara himpunan dapat ditulis $D = \{d_1, d_2, \ldots\ldots, d_i\}$. Kemudian struktur setiap dokumen disimbolkan dengan S dan setiap struktur disimbolkan dengan j maka ditulis S_j yang secara himpunan dapat ditulis $= \{s_1, s_2, \ldots\ldots, s_j\}$. Gabungan dokumen dan struktur ditulis dengan $Di\ x\ Sj$. Hasil dari perkalian sebagai fungsi F target T dapat ditulis $\phi: D\ x\ S \rightarrow \{T, F\}$. ϕ menggambarkan bagaimana dokumen diklasifikasi dengan struktur S. Penjelasan lebih detail di bagian ekperimen. Dalam menstrukturisasi dokumen diasumsikan sebagai Berikut :

- Kategori diasumsikan sebagai simbolik label sebagai bentuk format bukan sebagai knowledge tambahan yang berarti punya arti tertentu
- Bukan sebagai pengetahuan yang bersumber dari luar dokumen sehingga menimbulkan arti baru tapi strukturisasi ini sebagai penanda sebagai pengganti format.
-

2.2.Struktur Text sebagai Format dari Dokumen

Format dari setiap dokumen yang akan dianalisis direpresentasikan sebagai struktur yang akan dicari secara otomatis menggunakan metoda di machine learning. Beberapa riset untuk mengkategorikan struktur dari suatu text dari (Sebastiani, 2002) (Zinman *et al.*, 2006) (Wallach, 2008)(Rossi and Gnawali, 2014)(Noesgaard and Ørngreen, 2015).

2.3.Machine Learning

Machine Learning adalah metode untuk mempelajari data kemudian melakukan prediksi untuk masa berikutnya (Hu *et al.*, 2014). Machine learning terbagi atas 2 bagian yaitu bagian mesin learning supervised dan machine learning unsupervised. Machine Learning supervised adalah struktur dari suatu data yang hendak dianalisis telah ditentukan dahulu dan machine learning mencari data di sttuktur tersebut, sedangkan mesin learning unsupervised struktur dari suatu data dicari oleh machine learning itu sendiri(Pojon, 2017). Dalam riset ini machine learning yang digunakan yaitu unsupervised. Dipilih metode unsupervised karena struktur dari panduan skripsi adalah dinamis sehingga machine learning yang akan menentukan sendiri struktur dari data yang dimasukkan. Dokumen yang menjadi acuan model dari struktur format skripsi adalah panduan skripsi dari fakultas ilmu komputer. Dalam bentuk formalnya machine learning dengan metode naïve Bayes Classifier Berikut

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Dengan A dan B adalah dua document yang berbeda dan $P(A)\ dan\ P(B)$ adalah probabilitas dari struktur text.$P(A|B)$ kemungkinan dari

p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom

topik/struktur dari dokumen yang dianalisis. Dalam penelitian ini proses analisis dokumen menggunakan perangkat lunka Mallet. Mallet adalah perangkat lunak yang dapat digunakan untuk menganalisis dokumen dengan menggunakan metoda machine learning dengan teknik yang bisa digunakan seperti naïve bayes, conditional random Fields dan mengklasifikasi.

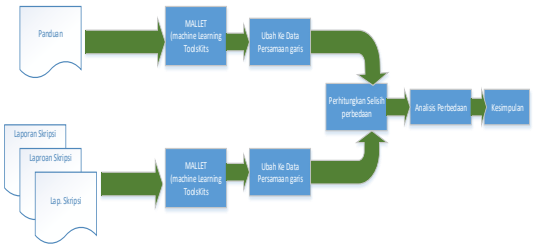
3. Metodology Penelitian

Metodology dalam penelitian ini sebagai Berikut :

1. Menentukan identtifikasi masalah
2. Membuat hipotesis
3. Menguji Hipotesis
4. Melakukan analisa dari hasil pengujian
5. Membuat kesimpulan

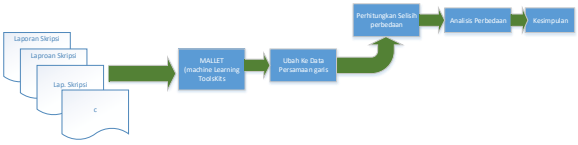
3.1.Metode Penelitian

Metode untuk mendeteksi kedekatan kesamaan pola. Pola yang didapat dari panduan skripsi yang kemudian dilakukan proses *match* atau penyamaan dengan pola dari skripsi mahasiswa. Untuk mengukur keakuratan dari pola dengan mengukur selisih pola persamaan garis antara laporan dan panduan. Berikut diagram di gambar 1. metode pengujian terpisah kesamaan format dokumen dengan membandingkan isi struktur yang terbentuk



Gambar 1 Metode Pengujian Dokumen skripsi terpisah dengan panduan skripsi

Untuk membuktikan bahwa kesimpulan dengan metode pengujian terpisah adalah akurat dibandingkan dengan metode pengujian dokumen dengan pembentukkan pola struktur yang disatukan. Berikut di gambar 2 metoda pengujian dokumen gabungan dengan panduan.



Gambar 2 Metode Pengujian Dokumen secara Gabungan

3.2. Identifikasi Masalah

Identifikasi masalah dari penelitian ini adalah

1. Apakah pembentukkan Pola struktur panduan skripsi dengan machine learning dapat digunakan untuk meneteksi kesesuaian format laporan skripsi.
2. Bagaimana cara menentukan perbedaan pola skripsi dan panduan untuk dapat menentukan akurasi skripsi yang dievaluasi.

3.3. Hipotesis

Hipotesis dari penelitian adalah dengan mengetahui pola panduan

p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom

skripsi maka dapat diukur akurasi kesesuaian pola skripsi mahasiswa.

3.1. Menguji Hipotesis

Dalam Melakukan pengujian hipotesis dilakukan langkah sebagai Berikut :

1. Membuat disain Eksperimen
2. Melakukan Eksperimen
3. Mencatat dan menganalisa hasil eksperimen

3.1.1. Disain Eksperimen

Perangkat Lunak dan Data yang digunakan.

Dalam Eksperimen ini perangkat lunak yang digunakan yaitu :

1. Perangkat Lunak mallet
2. Perangkat Lunak Microsoft Excell

Data yang digunakan untuk eksperimen :

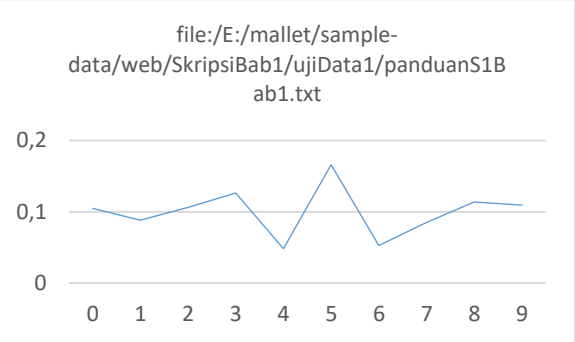
1. Panduan Skripsi Fakultas Ilmu Komputer
2. Skripsi mahasiswa

Langkah Eksperimen

1. Instalasi aplikasi mallet
2. Ubah file panduan skripsi BAB I ke bentuk Plain Text
3. Ubah file Skripsi Mahasiswa Khusus BAB I ke Plain Text
4. Sesuai dengan Panduan set Topik untuk analisis di mallet sesuai dengan panduan skripsi
5. Simpan semua dokumen dalam bentuk plain text ke lokasi folder yang akan dianalisis
6. Jalankan Proses analisis mallet untuk mendapatkan struktur dari panduan skripsi
7. Jalan proses analisis mallet digabung dengan dokumen skripsi yang formatnya sesuai dengan panduan
8. Jalankan proses analisis mallet digabung dengan dokumen skripsi yang format tidak sesuai
9. Proses diatas dijalankan secara terpisah untuk setiap mahasiswa
10. Ukur akurasi dengan menggunakan persamaan garis data panduan dan data skripsi yang diuji
11. Lakukan proses analisa dengan mallet secara gabungan untuk pembuktikan secara empiris bahwa kesimpulan secara terpisah atau gabungan adalah sama.

3.1.2. Eksperimen Pengecekan terpisah setiap mahasiswa.

Pengujian dengan mengggunakan mallet untuk membaca dokumen panduan skripsi BAB I didapat pola seperti Berikut :



Gambar 3 Pola struktur panduan Skripsi

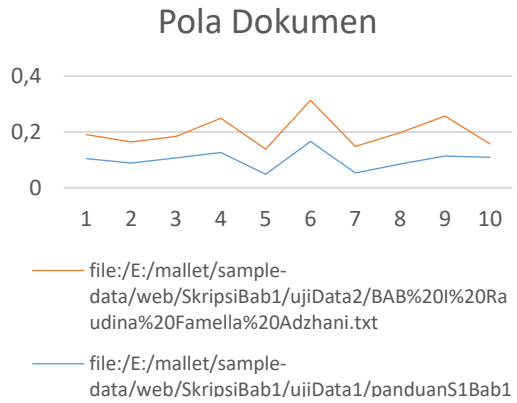
Dari gambar 3 angka di garis horizontal adalah jenis topik yang terbentuk sedangkan di garis vertical adalah persentasi isi dokumen terhadap struktur yang dibentuk. Bentuk data struktur tiap topik di tabel 1

Tabel 1 Tabel berisi topik yang dibentuk oleh machine learning mallet.

p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
0	0.5	dalam hal digunakan identifikasi data diidentifikasi hasil atas ada jadual laboratorium desa ditentukan dapayt tergantung lebih antara kamus diantaranya bergantung																		
1	0.5	hipotesis masalah dugaan harus berdasarkan apabila pernyataan perlu pemilihan apakah belakang perusahaan sekolah daerah alat terebut uji terstruktur sequential uang																		
2	0.5	dengan paragraf tujuan secara bagi teori kerangka variabel diuji kepada tepat penelitiannya diambil juga lembaga keacakan melalui perkataan benar-benar hasudah																		
3	0.5	penelitian dilakukan sampel ditumuskan aspek diteliti sehingga mengembangkan statistik batasan apa itu dua kota hendalnya sesungguhnya randomisasi tidalnya adanya suat																		
4	0.5	didasarkan bentuk tentang produk masalah menarik pendorong skripsi ujinya keperluan dikenal jadi faktor ilmu fokus sistem perbedaan menjawab dicapai umum																		
5	0.5	yang dari tidak untuk manfaat lain dinyatakan sedangkan maksud pertanyaan-pertanyaan biasanya disamping tersebut mewakili jawaban didefinisikan membangun linier prototy																		
6	0.5	untuk pendekatan manfaat diteliti perangkat metode menjelaskan lamanya waktu lokasi oleh didukung menggunakan singkat subbab dahulu tertentu bukan dasarnya berkaitan																		
7	0.5	sebagai adalah proses permasalahan dipilih teoritis berisi mungkin bahwa latar menguraikan jelas diperhatikan hanya hubungan dirancang langkah lapangan ditrunkan kemudia																		
8	0.5	dapat atau pada akan pemikiran sistem diteliti penelitimaka pengembangan dimana kerangka fenomena penjelasan spesifik kaidah-kaidah dfd program terlebih baku																		
9	0.5	ini dan model bila populasi objek perangkat dibuat berisikan sebagainya serta umi event berjalan menggambarkan variabel urutan lazim sampel dedukti																		

Kemudian diambil skripsi salah satu mahasiswa yang bernama Rd dan dibandingkan bentuk pola dengan panduan. kemudian dianalisis polanya dengan mallet didapatkan bentuk di gambar 4. Di gambar 4 tampak 2 grafik yaitu grafik panduan bab 1 dan grafik bab 1 milik Rd. Bila dilihat sekilas tampak bahwa format yang sudah dibuat oleh Rd hampir menyerupai dengan panduan. Bila diperhitungkan secara presisi maka perlu menggunakan persamaan garis untuk dapat mengetahui keserupaan pola. Untuk menghitung persamaan garis dapat digunakan formula



Gambar 4 Pola dokumen mahasiswa dibandingkan dengan panduan

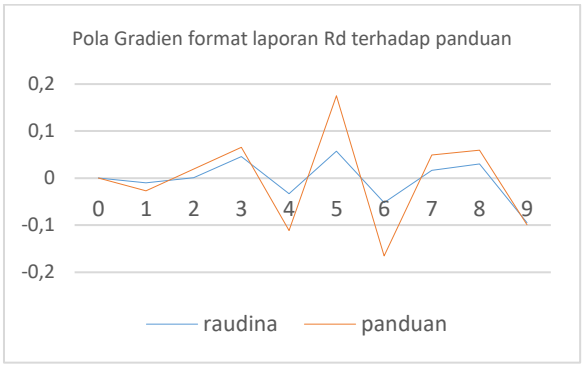
$$m = \frac{y2 - y1}{x2 - x1}$$

Dengan m adalah nilai gradient maka didapat seperti di tabel 2.

Tabel 2 nilai gradient yang didapat

	0	1	2	3	4	5	6	7	8	9
raudina	0	-1.01%	0.14%	4.59%	-3.31%	5.74%	-5.20%	1.69%	3.04%	-9.46%
panduan	0	-1.68%	1.83%	1.99%	-7.80%	11.77%	-11.31%	3.21%	2.91%	-0.46%

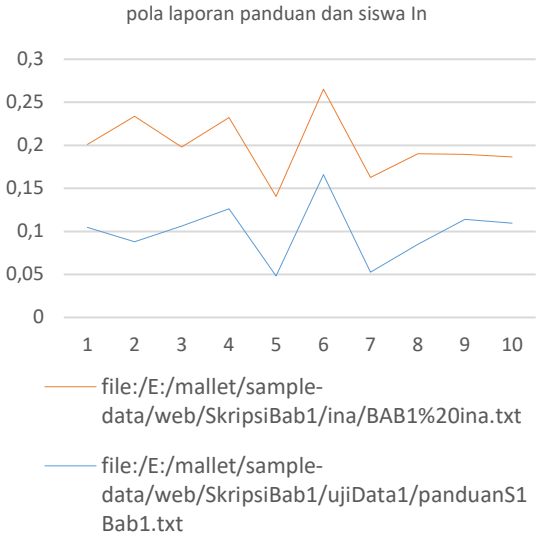
Bentuk grafik dari data tampak di gambar 5.



Gambar 5 Pola Gradien Format laporan Rd terhadap Panduan.

Di gambar 5 terlihat bahwa pola persamaan garis siswa Rd dengan panduan hampir menyerupai walaupun tidak sama persis. Perbedaan persentasi Topik panduan dan Rd diatas dirata rataka dengan nilai absolut (nilai negative dipositifkan) didapat perbedaan sebesar 3.58 %. Bila ditetapkan oleh fakultas bahwa perbedaan format dan panduan tidak boleh lebih dari 5 % maka laporan yang dibuat oleh Rd sudah memenuhi syarat karena nilai rata rata kurang dari maksimum perbedaan yang ditetapkan.
Percobaaan Berikut diambil data mahasiswa yang lain untuk dianalisa pola laporannya.
Dihitung nilai m didapat tampak di tabel 3, dan bila digambarkan secara grafik tampak di gambar 7. Tampak di gambar bahwa pola laporan siswa yang bernama In tidak sesuai panduan. Bila diperhitungkan nilai rata rata perbedaan graiden sebesar 6,05 % dan ini tidak memenuhi syarat ketetapan fakultas karena diatas 5 %.

p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom



Gambar 6 Pola dokumen mahasiswa dibandingkan dengan panduan

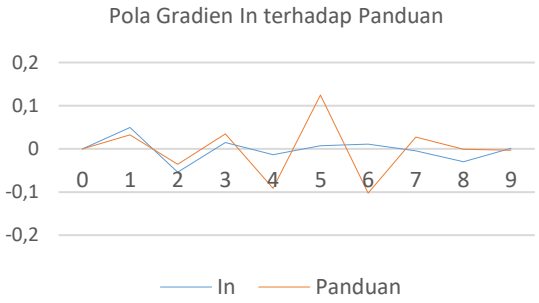
3.1.3. Eksperimen pengecekan Gabungan semua mahasiswa

Eksperimen gabungan Berikut untuk menguatkan kesimpulan yang didapatkan dari eksperimen terpisah yang telah dilakukan di sub 3.1.2 bahwa kesimpulan laporan siswa yang bernama Rd lebih baik dari siswa bernama In. Eksperimen ini melakukan analisa panduan dan laporan mahasiswa untuk membentuk struktur secara

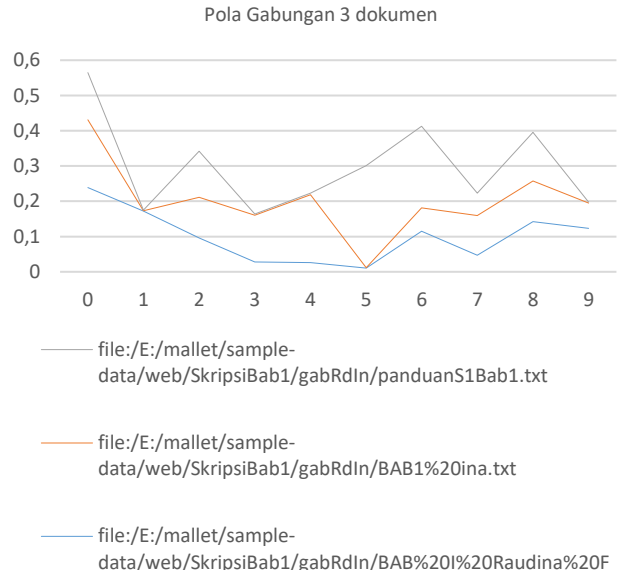
Tabel 3 Nilai persamaan garis mahasiswa bernama In

In	0.00%	4.95%	-5.40%	1.45%	-1.37%	0.68%	1.07%	-0.46%	-2.97%	0.15%
Panduan	0	-1.68%	1.83%	1.99%	-7.80%	11.77%	-11.31%	3.21%	2.91%	-0.46%

bersamaan sehingga dapat dibentuk struktur topik yang berlaku untuk ketiga dokumen seperti tampak di gambar 7.
Tampak di gambar 7 bahwa pola laporan siswa bernama Rd lebih baik dari siswa bernama In, dan ini bisa dibuktikan dengan menghitung perbedaan gradient dari Rd dan In terhadap panduan didapatkan Rd sebesar 11 % dan In sebesar 14 %.



Gambar 7 Pola Gradien format laporan mahasiswa dibandingkan dengan panduan



Gambar 8 Pola dokumen dengan pembentukkan struktur secara bersamaan.

4. Kesimpulan

1. Berdasarkan percobaan yang telah dilakukan membuktikan bahwa Machine Learning mampu mendeteksi kesesuaian laporan skripsi mahasiswa terhadap panduan penyusunan laporan
2. Mendeteksi kesesuaian laporan dapat dilakukan dengan cara terpisah yaitu masing masing dokumen dianalisis polanya dengan machine learning untuk mendapatkan pola struktur

p-ISSN : 1858-3911, e-ISSN : 2614-5405
https://journal.uniku.ac.id/index.php/ilkom

dokumen kemudian dilakukan komparasi menggunakan metoda persamaan garis

3. Mendeteksi kesesuaian laporan dapat juga dilakukan dengan cara menganalisis secara gabungan seluruh dokumen terhadap panduan untuk pola struktur yang terbentuk. Kemudian dianalisis menggunakan metoda persamaan garis.

5. Penelitian ke Depan

Dari penelitian yang telah dibahas menunjukkan bahwa machine learning dapat digunakan untuk mendeteksi besarnya persentase kesesuaian panduan dan format laporan di bab 1. Yang belum diteliti adalah mendeteksi kesesuaian margin dan spasi dari laporan penelitian. Dan perlu diperhatikan juga dengan keterbatasan dari machine learning (Hoos, 2017) , sehingga perlu digabungkan dengan metode lain untuk meningkatkan akurasi dari machine learning.

Referensi

- Hoos, H. H. (2017) 'Machine Learning – Opportunities and Limitations'.
- Hu, {yuening *et al.* (2014) 'Interactive Topic Modeling', pp. 0–56. Available at: http://cs.colorado.edu/.
- Noesgaard, S. S. and Ørngreen, R. (2015) 'The effectiveness of e-learning: An explorative and integrative review of the definitions, methodologies and factors that promote e-Learning effectiveness', *Electronic Journal of e-Learning*, 13(4), pp. 278–290.
- Pojon, M. (2017) 'Using Machine Learning to Predict Student Performance', *India*, (June).
- Rossi, L. A. and Gnawali, O. (2014) 'Language independent analysis and classification of discussion threads in Coursera MOOC forums', *Proceedings of the 2014 IEEE 15th International Conference on Information Reuse and Integration, IEEE IRI 2014*, pp. 654–661. doi: 10.1109/IRI.2014.7051952.
- Sebastiani, F. (2002) 'Machine Learning in Automated Text Categorization', 34(1), pp. 1–47.
- Wallach, H. M. (2008) 'Structured Topic Models for Language', *Doctor*, (2001), p. 136. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.149.2537&rep=rep1&type=pdf.
- Wuillemin, P.-H. and Torti, L. (2012) 'Structured probabilistic inference', *International Journal of Approximate Reasoning*. Elsevier, 53(7), pp. 946–968. doi: 10.1016/J.IJAR.2012.04.004.
- Zinman, A. *et al.* (2006) 'Probabilistic Topic Models', *MIS Quarterly*, 3(3), pp. 993–1022. doi: 10.1016/s0364-0213(01)00040-4.