

Berkala Ilmu Perpustakaan dan Informasi, Vol. 16, No. 2, Desember 2020, Hal. 142-154
DOI: 10.22146/bip.v16i1.424
ISSN 1693-7740 (Print), ISSN 2477-0361 (Online)
Tersedia online di <https://journal.ugm.ac.id/v3/BIP>

Identifikasi kecakapan inovasi lembaga riset di Indonesia berbasis dokumen

Aris Yaman¹, Bagus Sartono², Agus M. Soleh³

¹Pascasarjana, IPB University

^{2,3}Departemen Statistika dan Sain Data, IPB University

e-mail: arisyaman@apps.ipb.ac.id

Naskah diterima: 22 Juni 2020, direvisi: 7 Oktober 2020, disetujui: 19 Oktober 2020

ABSTRAK

Pendahuluan. Inovasi yang dihasilkan lembaga-lembaga riset di Indonesia terindikasi adanya duplikasi. Pemetaan spesialisasi inovasi lembaga riset penting untuk dilakukan. Penelitian ini mengkhususkan pada pemetaan kecakapan inovasi lembaga riset di Indonesia.

Metode penelitian. Penelitian ini menggunakan metode analisis dokumen teknologi berbasis paten untuk memetakan potensi teknologi. Data yang digunakan merupakan data paten terdaftar pada database DJKI.

Data analisis. Metadata analisis dilakukan menggunakan metode *K-means Clustering* dengan bantuan *software R*.

Hasil dan Pembahasan. Temuan pada *pre* analisis menunjukkan bahwa ketika peubah bebas yang terlibat dalam model sangat besar, metode *Localized feature selection*, dapat efektif menseleksi peubah tanpa kehilangan banyak informasi. Hasil analisis menggunakan metadata paten menunjukkan bahwa terdapat 5 kelompok teknologi dominan yang mampu dihasilkan oleh lembaga riset di Indonesia, antara 1) Teknologi terkait pengembangan teknologi instrumen pengukuran dan pengujian; 2) Teknologi terkait makanan dan bahan makanan; dan 3) Alat uji mikrostruktural; 4) Radar dan sejenisnya; 5) Teknologi dalam bidang pertanian.

Kesimpulan dan Saran. Temuan menunjukkan bahwa masih terdapat tumpang tindih inovasi oleh beberapa lembaga riset yang berbeda pada kluster teknologi yang sama. Selain itu, *K-means Clustering* dengan *pre* analisis LFSBSS memiliki performa yang baik dalam memisahkan kluster teknologi dengan jelas.

Kata kunci: k-means clustering; paten; teknologi spesialis; metadata analisis

ABSTRACT

Introduction. Duplication in inventions produced by research institutions in Indonesia becomes an issue. It is important to map the specialization of the invention in research institutions. This study examines the mapping of the innovation in research institutions in Indonesia.

Data Collection Method. This study uses a patent-based technology document analysis method to map the potential of technology. The data used is patent data registered in the Direktorat Jenderal Kekayaan Intelektual (DJKI) database.

Data Analysis. Metadata analysis was conducted by using the *K-Means Clustering* method with *R* software.

Results and Discussions. The findings in the *pre*-analysis show that when the independent variable involved in the model are very large, the *Localized feature selection* method can effectively select variables without losing much information. There are 5 dominant technology groups that can be produced by research institutions in Indonesia, namely 1) Technology related to the development of measurement and testing instrument technology; 2) Technologies related to food and food ingredients; and 3) microstructural test equipment / detectors; 4) radar technology; 5) Technology in agriculture.

Conclusion. *The findings show that there are still overlapping inventions by several research institutions in the same technology cluster. K-means clustering with LFSBSS pre analysis has a clear performance in the technology cluster space.*

Keywords: *k-means clustering; patent; technology specialist; metadata analysis*

A. PENDAHULUAN

World Economic Forum (WEF) 2019 menuturkan bahwa daya saing bangsa Indonesia mengalami penurunan lima level dari periode sebelumnya, dari posisi 45 ke posisi 50 di periode 2018-2019. Salah satu faktor yang menyebabkan rendahnya daya saing Indonesia adalah rendahnya kapasitas bangsa dalam berinovasi, indeks poin dalam kategori ini termasuk rendah hanya 37.7 dalam 100 poin maksimal (Schwab, 2019).

Pemangku kebijakan (Kemenristek/BRIN) kiranya dapat memetakan kecakapan inovasi lembaga-lembaga riset dalam otoritasnya. Saat ini setidaknya ada empat lembaga riset besar dalam koordinasi BRIN yaitu LIPI, BPPT, LAPAN dan BATAN (Presiden RI, 2019). Pemetaan kecakapan inovasi ini penting dilakukan supaya pemangku kepentingan dapat mengetahui arah kebijakan teknologi yang sesuai dengan kapasitas negaranya. Menindaklanjuti adanya tumpang tindih output antar lembaga riset.

Penelitian dan pengembangan suatu lembaga riset dapat ditelusuri melalui dokumen paten yang terdaftar di kantor paten. Paten merupakan hak eksklusif yang diberikan oleh negara kepada inventor atas hasil penemuannya di bidang teknologi, yang selama waktu tertentu melaksanakan sendiri penemuan tersebut atau memberikan persetujuan kepada pihak lain untuk melaksanakannya. Hak paten merupakan salah satu bentuk dari Hak Kekayaan Intelektual (HKI). Paten dapat meningkatkan profitabilitas dan mempertahankan keunggulan kompetitif (Artz et al., 2010).

Dokumen-dokumen paten berisi berbagai informasi teknis yang berkaitan dengan hak kekayaan intelektual dan hasil penelitian (Chen et al., 2017). Sebuah penelitian menunjukkan bahwa negara yang tidak melakukan analisis terhadap data paten berdampak terhadap

tingginya biaya dan waktu yang panjang dalam proses pengembangan teknologi (Kim & Bae, 2017). Oleh karena itu, pemangku kebijakan seharusnya dapat memanfaatkan informasi yang terdapat pada penelitian ilmiah dan pengembangan industri yang telah dipatenkan sebelumnya untuk menentukan arah inovasi bangsa. Dalam hal ini, jika digunakan secara benar, informasi paten dapat menjadi salah satu faktor pendorong utama dalam perkembangan teknologi dan kemajuan sosial. Pemanfaatan data paten dapat dijadikan sebagai landasan dalam menyusun roadmap inovasi. Hal ini tidak hanya memainkan peran penting dalam menemukan peluang teknologi baru, menghindari duplikasi investasi R&D, mencegah risiko pelanggaran paten, dan sebagainya, tetapi juga memiliki arti penting untuk memandu pengembangan industri secara keseluruhan (Yu & Zhang, 2019).

Peneliti-peneliti terdahulu telah melakukan penelitian terkait isu-isu paten yang berkembang dari berbagai perspektif. Penelitian paten sebelumnya pada umumnya menggunakan pendekatan metode bibliometrika dalam menganalisis secara kuantitatif literatur ilmiah (B. Wang et al., 2018). Su et al. (2019) menggunakan metode terintegrasi *main path analysis* untuk menganalisis jalur arah pengembangan di bidang ilmu pengetahuan dan teknologi.

Analisis paten membantu dalam menganalisis kronologi status teknologi saat ini, peramalan tren teknologi, pertumbuhan ekonomi, kapasitas teknologi nasional, daya saing, nilai pasar, kemampuan R&D, dan perencanaan strategis untuk menghindari pengeluaran R&D yang tidak perlu (Y. L. Wang, 2012).

Berbagai penelitian terdahulu yang mengidentifikasi teknologi berbasis dokumen paten. Analisis lebih banyak dilakukan berbasis analisis kata kunci (*text mining*). Padahal sebuah

dokumen paten sudah secara global disepakati untuk dicantumkan kode teknologi berdasarkan suatu pedoman tertentu. Pengkodean teknologi pada dokumen paten ini dikenal dengan istilah IPC patent (WIPO, 2018). Diharapkan analisis dengan menggunakan IPC paten lebih konsisten dan dengan tingkat validitas yang lebih baik.

Kajian ini berusaha mengisi gap riset yang terjadi ketika menganalisis hanya berbasis analisis kata kunci. Penelitian ini memfokuskan analisis pada kode IPC yang terdapat dalam dokumen paten.

Mengambil paten yang sesuai dari dataset besar ini adalah suatu yang kompleks. Proses ini dapat disederhanakan jika dapat membagi dataset menjadi kelompok. Analisis kluster merupakan salah satu solusi yang dapat melakukan hal ini (Dhir et al., 2019).

Salah satu aspek yang dapat menimbulkan ambiguitas pada analisis kluster terjadi ketika peubah yang dianalisis cukup banyak. Hal ini berdampak pada sulitnya intepretasi kluster yang dihasilkan nantinya. Penelitian ini menambahkan pre analisis dengan melakukan seleksi peubah terlebih dahulu sebelum dilakukan analisis kluster. Hasil kluster dengan metode ini lebih mudah untuk ditafsirkan dan intepretasikan.

Penelitian ini mengidentifikasi teknologi utama pada bidang tertentu dan menemukan bagaimana teknologi dari bidang yang berbeda dapat saling terkait dan terintegrasi. Analisis klasifikasi paten dapat digunakan untuk memetakan alur perkembangan teknologi, masa lalu, masa kini, dan prediksi masa datang, bahkan jauh sebelum produk muncul di pasar.

Tujuan penelitian ini adalah memetakan kecapan inovasi lembaga riset di Indonesia berbasis paten kluster, menggunakan database paten DJKI Kemenkumham (Direktorat Jenderal Kekayaan Intelektual, Kementerian Hukum dan HAM) melalui analisis IPC (*International Patent Classification*), sehingga didapatkan area strategis untuk pengembangan inovasi.

B. TINJAUAN PUSTAKA

International Patent Classification (IPC)

International Patent Classification (IPC) merupakan suatu sistem klasifikasi paten secara internasional yang dikembangkan oleh *World Intellectual Property Organization (WIPO)*. Sistem klasifikasi IPC merupakan cara yang paling tepat untuk menggali informasi paten karena pengkategorian dan pengindeksannya yang konsisten. Untuk mengklasifikasikan semua peubah bebas teknologi yang relevan, seluruh paten diklasifikasikan setidaknya oleh satu kode IPC. Ketika terdapat lebih dari satu aspek teknologi yang berbeda dalam satu paten tunggal, maka paten tersebut akan memiliki lebih dari satu kode IPC. Sebagai hasilnya, dapat dengan mudah untuk diidentifikasi berapa banyak teknologi dan bidang teknologi mana yang saling terkait dalam suatu paten. Aplikasi paten di masing-masing bidang menunjukkan akumulasi dari pengetahuan dan kemajuan dalam lintasan teknologi (*Fai & von Tunzelmann, 2001*). Oleh karena itu, analisis data kode IPC yang diekstraksi dari dokumen paten memungkinkan untuk memahami perkembangan teknologi dan membuat prediksi tren pengembangan teknologi ke depan.

Kode IPC adalah suatu hirarki yang menetapkan keberadaan paten dalam suatu kategori. Terdapat 8 Seksi, 131 kelas, 642 subkelas, dan 73.915 grup. Adapun 8 Seksi IPC terdiri dari (A) Kebutuhan Manusia (*Human Necessities*); (B) Pengoperasian, Transportasi (*Perfoming Operation, Transportation*); (C) Kimia, Metalurgi (*Chemistry, Metallurgy*); (D) Tekstil, Kertas (*Textiles, Paper*); (E) Konstruksi (*Fixed Constructions*); (F) Teknik Permesinan, Penerangan, Pemanasan, Senjata, Peledakan (*Mechanical Engineering, Lighting, Heating, Weapons, Blasting*); (G) Fisika (*Phyisics*); (H) Kelistrikan (*Electricity*). Peneliti menggunakan versi IPC (2018) dalam penelitian ini (WIPO, 2018). Sebagai contoh, definisi dari A61 adalah “Ilmu Medis dan Kedokteran Hewan; Ilmu Kesehatan” dibagi dalam 16 subkelas, antara lain, A61B (Diagnosa, Operasi, dan Identifikasi), A61D (Instrumen, Pengaplikasian, Alat, dan Metode untuk

Kedokteran Hewan), dan A61K (Persiapan untuk tujuan medis, gigi, atau toilet).

K-Means Clustering

Dokumen paten dari berbagai invensi merupakan data yang sangat besar dan kompleks. Menjadi rumit jika kita menganalisis dengan metode statistika deskriptif. Proses ini dapat disederhanakan jika dataset dapat dibagi menjadi kelompok. Analisis kluster merupakan solusi yang dapat melakukan hal ini (Dhir et al., 2019). Analisis kluster salah satu metode deskriptif statistik yang bertujuan mengklusterkan objek-objek menjadi beberapa kelompok/kluster, berdasarkan ukuran kemiripan/ketakhiripan antar objek. Berbagai metode digunakan untuk menentukan keanggotaan kelompok dalam analisis kluster. Secara umum analisis kluster dibagi menjadi dua yaitu Analisis Kluster Hierarki dan Analisis Kluster Non-Hierarki.

Salah satu metode analisis kluster non hierarki atau metode *partitioning* yang paling populer adalah Analisis kluster K-means. Pada metode ini banyaknya kluster secara pasti harus sudah diketahui. Sebuah cara sederhana yang dapat digunakan untuk menentukan banyak kluster yang akan digunakan adalah plot dari jumlah kuadrat dalam kluster (*within sum of squares*) dengan banyak kluster yang mana hasilnya mirip dengan *screeplot* dalam analisis faktor.

Localized Feature Selection Based on Scatter Separability (LFSBSS)

Fenomena yang sering terjadi dalam analisis kluster diantaranya terdapat peubah-peubah bebas yang tidak relevan dalam kumpulan data yang dapat mengurangi kualitas akhir pengklusteran dan menghabiskan lebih banyak memori. Oleh karena jika peubah bebas ini dihapus, waktu komputasi dapat dihemat. Dari sudut pandang pengelompokan, menghapus peubah bebas yang tidak relevan tidak akan berdampak negatif pada akurasi pengelompokan sekaligus mengurangi penyimpanan dan waktu komputasi yang diperlukan (Aggarwal, Charu C.; Reddy, 2014).

Ide LFSBSS didasari fenomena bahwa peubah yang relevan dengan satu hasil pengklusteran tidak perlu set yang sama yang relevan dengan hasil pengklusteran lain (Li et al., 2008). Dengan kata lain suatu peubah bebas hanya relevan bagi satu atau beberapa kluster yang terbentuk, tidak untuk semua kluster. Li et al. (2008) dalam mengevaluasi kriteria peubah bebas yang relevan dalam suatu kluster, mengadopsi rumusan plot separabilitas yang telah ditemukan sebelumnya.

LFSBSS mengurangi adanya dampak tumpang tindih dan tidak terkategori-sasinya suatu amatan dengan memberikan fungsi penalti. Fungsi ini disebut sebagai normalisasi nilai yang dijustifikasi. Nilai proyeksi dengan penalti ini diterapkan jika jumlah amatan yang saling tumpang tindih atau tidak terkategori semakin meningkat dibandingkan dengan nilai proyeksi pengklusteran asli/awal. LFSBSS pada prinsipnya mengadopsi *backward feature selection*. Hal ini berarti bahwa kluster dibuat terlebih dahulu menggunakan seluruh peubah penjelas, kemudian, secara berulang menghapus peubah penjelas yang tidak relevan (Aggarwal, Charu C.; Reddy, 2014).

Principal Component Analysis (PCA) atau Analisis Komponen Utama (AKU)

Analisis PCA ini pada dasarnya bertujuan untuk menghilangkan adanya korelasi antar peubah penjelas. PCA memberikan informasi yang cukup dengan sejumlah peubah baru yang lebih sedikit yang saling tegak lurus yang disusun dari kombinasi peubah asal. Dengan kata lain diusahakan didapatkan sejumlah gugus peubah baru yang diolah dari sejumlah peubah lama sedemikian sehingga menghasilkan gugus peubah baru. Gugus baru peubah ini jauh lebih sedikit dibanding banyaknya gugus peubah lama. Peubah baru ini merupakan kombinasi linear dari gugus peubah lama. Peubah-peubah baru yang terbentuk memiliki sifat saling bebas (Markowitz, 2018).

Atas dasar bahwa peubah baru yang terbentuk dari proses PCA ini jauh lebih sedikit dibandingkan jumlah peubah sebelumnya, maka hasil perhitungan peubah baru ini dapat

diadaptasi untuk analisis lanjutan. Beberapa analisis lanjutan yang memungkinkan yaitu analisis regresi PCA dan Kluster PCA.

Secara matematis tujuan PCA pada dasarnya adalah memutar sumbu ruang p-dimensi secara kaku ke posisi baru (sumbu utama) yang memiliki sifat yaitu sedemikian rupa sehingga sumbu utama 1 memiliki varian tertinggi, sumbu 2 memiliki varian tertinggi berikutnya, dan sumbu p memiliki varian terendah. Kovariansi di antara setiap pasang sumbu utama adalah nol (sumbu utama tidak berkorelasi) (Johnson & Wichern, 2007).

Secara konsep teori statistika, pembentukan komponen utama pada prinsipnya merupakan proses terbentuknya kombinasi linear dari peubah-peubah penjelas yang diamati. Oleh karena kombinasi linear tentunya diperlukan suatu metode yang dapat memberikan bobot dari tiap peubah penjelas yang diamati.

Proses mendapatkan nilai-nilai koefisien atau bobot dari kombinasi linier peubah-peubah penjelasnya memiliki ketentuan sebagai berikut, pertama ada sebanyak p komponen utama, yaitu sebanyak peubah yang diamati dan setiap komponen utama adalah kombinasi linier dari peubah-peubah tersebut. Kedua, setiap komponen utama saling ortogonal (tegak lurus) dan saling bebas. Terakhir, komponen utama dibentuk berdasarkan urutan variansi dari yang terbesar hingga yang terkecil (Zelterman, 2015).

C. METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan metadata dokumen paten yang terdapat dalam database paten DJKI Kemenkumham (Direktorat Jenderal Kekayaan Intelektual, Kementerian Hukum dan HAM), dengan alamat <https://pdki-indonesia.dgip.go.id/>. kata kunci dalam pencarian yang digunakan: LIPI, BPPT, BATAN dan LAPAN. Terdapat sebanyak 280 observasi (dokumen paten) yang tersebar pada lima lembaga riset tersebut. Metadata dari 280 dokumen paten ini diklasetrisasi dengan metode *K-Means Clustering*.

DJKI merupakan satu-satunya insititusi di Indonesia yang diberikan mandat oleh pemerintah dalam mengelola dan

memverifikasi paten (Presiden Republik Indonesia, 2016). LIPI, BPPT, BATAN dan LAPAN merupakan lembaga riset terbesar di Indonesia sehingga dipilih sebagai objek dalam penelitian ini.

Analisis lebih lanjut dilakukan khususnya terhadap kode IPC dalam metadata dokumen paten. Penyampaian metode analisis yang dilakukan dalam kajian ini digambarkan dalam bentuk diagram alir seperti tampak pada Gambar 1.

Diagram alir penelitian yang terdapat dalam Gambar 3 dapat dinarasikan sebagai berikut:

1. Identifikasi banyaknya peubah penjelas yang akan diteliti dalam mencari fenomena tertentu
2. Identifikasi apakah jumlah peubah yang dianalisis banyak dan atau melebihi banyaknya amatan. Hal ini penting dilakukan karena apabila terjadi jumlah amatan lebih sedikit dibandingkan jumlah peubah penjelas menyebabkan terbentuknya matriks singular. Hal ini berdampak tidak akan didapat solusi penyelesaian.
3. Jika terjadi fenomena pada langkah 2, maka perlu dilakukan pre analisis, peneliti menerapkan pre analisis dengan PCA. Jika prosedur PCA masih menimbulkan multitafsir ketika dilakukan analisis lanjutan dengan *K-means Clustering*, maka perlu dilakukan pre analisis dengan pendekatan LFSBSS.
4. Jika dirasa pre analisis dengan PCA dan atau LFSBSS cukup informatif, kita lakukan pengklasteran berdasarkan hasil seleksi peubah yang telah dilakukan pada tahap PCA dan atau LFSBSS.
5. Melakukan analisis klaster dengan metode *K-Means Clustering*. Secara ringkas metode ini dapat dijelaskan sebagai berikut:
 - a. Masukan sebanyak 280 dataset paten kedalam algoritma k-means. Tentukan sejumlah p peubah bebas hasil seleksi peubah, baik dengan metode PCA dan atau dengan LFSBSS
 - b. Tentukan banyaknya klaster yang ingin dibentuk (penelitian ini dilakukan

- dengan pendekatan screeplot)
 - c. Tentukan ukuran pemusatan kluster (*centroid*) untuk setiap kluster yang yang ditentukan pada tahap sebelumnya
 - d. Ukur jarak setiap observasi dengan centroid pada tahap sebelumnya
 - e. Kelompokkan setiap observasi berdasarkan jarak terkecil
 - f. Identifikasi apakah observasi berpindah kluster
 - g. Jika tidak terjadi lagi pemidahan posisi observasi, kluster yang terbentuk sudah valid, jika masih terdapat perpindahan observasi ke kluster yang lain, ulangi dimulai dari langkah d).
6. Intrepretasi hasil.

D. HASIL DAN PEMBAHASAN

Distribusi Kecakapan Inovasi Lembaga Riset Indonesia

LIPI menjadi lembaga riset yang memiliki tingkat produktivitas paten tertinggi dibanding tiga lembaga litbang lainnya. Lima Puluh satu persen dari keseluruhan paten lemlitbang diproduksi oleh LIPI, diikuti Batan dengan perolehan 24% paten nasional, kemudian BPPT sebesar 21%, dan yang masih dalam tahap perkembangan yaitu LAPAN sekitar 4% dari total paten lembaga litbang (Gambar 2). Hal ini mengindikasikan bahwa LIPI mempunyai SDM yang cukup cakap dalam berbagai kategori teknologi dan paling produktif dalam menghasilkan paten dibandingkan tiga lembaga litbang lainnya.

Gambar 2 (bagian distribusi berdasarkan kategori teknologi), memperlihatkan distribusi kategori inovasi lembaga RnD di Indonesia. Gambar tersebut memperlihatkan secara global, kategori invensi lembaga litbang di Indonesia didominasi oleh invensi dengan IPC tipe A61 (ilmu kesehatan atau veteriner), A23 (makanan atau bahan makanan) dan C12 (biokimia; mikrobiologi; enzimologi; mutasi atau teknik genetic), dengan persen kontribusi terurut sebesar 13%, 9% dan 8%. Hal ini sejalan dengan perkembangan industri farmasi di tanah air. Sebesar 73% pangsa pasar farmasi nasional didominasi oleh perusahaan farmasi lokal

(Mubarak, 2017).

Gambar 4, diagram kue bagian kategori invensi juga memberikan suatu indikasi bahwa 3 teknologi dominan atau teknologi yang paling banyak dikembangkan di Indonesia hanya sebesar 30%. Sedangkan sebanyak 70% teknologi lainnya menyebar beragam diberbagai kategori teknologi. Hal ini mengindikasikan lembaga-lembaga riset pada empat lembaga ini belum memiliki fokus teknologi yang jelas. Dengan kondisi sumber daya dan ketertinggalan riset Indonesia saat ini, sulit rasanya jika kita dapat mengembangkan teknologi yang beragam dengan maksimal. Hal ini menjadi gambaran perlunya memusatkan kembali fokus pengembangan suatu teknologi di Indonesia. Dalam menilai dan atau menentukan pusat kecakapan teknologi per lembaga dianalisis lebih lanjut dengan metode *K-means Clustering* dengan pre analisis LFSBSS pada analisis berikutnya.

Pengklasteran Kecakapan Inovasi Nasional

Dalam kasus ini terdapat sejumlah 69 peubah bebas (contoh H01, A23, A01, A01, C08, C08, B27, B27, C11, C08, G01, B27, G10, B27, B81, G06) (penjelasan kode lihat WIPO, 2018) dengan 280 observasi yang akan dianalisis lebih lanjut. Analisis lanjut yang digunakan dalam kajian ini menggunakan metode *K-means Clustering*.

Sebelum melakukan analisis kluster *kmeans*, diasumsikan bahwa jumlah peubah bebas yang dianalisis termasuk besar, sehingga perlu dilakukan pre analisis terlebih dahulu untuk mereduksi jumlah peubah bebas tanpa kehilangan yang berarti persen keragaman model (perhatikan diagram alur analisis pada Gambar 3).

Setelah dilakukan pre analisis dengan LFSBSS, diketahui bahwa dari 69 peubah bebas terseleksi sebanyak 8 peubah bebas yang memiliki tingkat kontribusi yang signifikan terhadap model (table 1). Peubah bebas hasil seleksi dianalisis lebih lanjut yaitu H01, A23, A01, B27, C11, G01, G10, B81. Dengan sejumlah delapan peubah bebas diharapkan memudahkan interpretasi kluster. Apabila kita bandingkan analisis pada Gambar 2 dan Tabel 1,

terlihat bahwa kategori teknologi dominan A61 dan C12 tidak masuk dalam filter dengan LFSBSS. Hal ini mengindikasikan bahwa teknologi kategori A61 dan C12 bukan merupakan peubah pembeda atau peubah pemisah yang dapat memisahkan dengan jelas antar kluster teknologi yang terbentuk. Berdasarkan hal ini, penelitian-penelitian sebelumnya yang berbasis bibliometrik dengan analisis hanya mendasarkan pada nilai paling tinggi/paling banyak muncul pada kategori tertentu, menjadi riskan untuk diambil sebagai dasar kebijakan. Hal ini karena dalam metadata yang dianalisis pada penelitian ini ditemukan fenomena, teknologi yang dominan/paling besar belum tentu merupakan teknologi pembeda yang menjadi dasar pusat kecakapan teknologi. Bahkan apabila kita elaborasi lebih lanjut, A61 dan C12 ini merupakan kategori-kategori teknologi yang diteliti oleh banyak institusi lembaga riset. Dengan kata lain indikasi tumpang tindih riset terjadi pada kategori teknologi ini. Hasil LFSBSS pada Tabel 1. juga memberikan gambaran teoritis bahwa metode ini dapat dengan jelas dan baik memisahkan mana peubah yang dapat menjadi peubah pemisah dalam pembentukan kluster nantinya.

Delapan peubah bebas ini dianalisis lebih lanjut dengan prosedur *k-means clustering*. Penentuan jumlah kluster efektif didasarkan pada *scree plot* antara banyaknya kluster dan jumlah kuadrat dalam kluster. Jumlah kluster efektif didapati ketika penambahan jumlah kluster tidak lagi menambah jumlah kuadrat dalam kluster secara signifikan. Secara grafik hal ini terlihat ketika kurva pada *screeplot* mulai melandai. Pada kondisi ini jumlah kluster terbaik didapatkan. Gambar 3. memperlihatkan secara visual kurva mulai melandai disaat jumlah kluster sebanyak tujuh. Oleh karena itu, ditentukan jumlah kluster sebesar tujuh ($k=7$) dalam analisis *K-means Clustering*.

Setelah didapat jumlah kluster optimum sebanyak tujuh, dilakukan analisis kluster *kmeans*. Hasil analisis kluster dengan metode *kmeans* dapat dilihat pada Tabel 2. Data pada Tabel 2, menampilkan hasil analisis kluster terhadap metadata dokumen paten pada lembaga-lembaga riset di Indonesia. Nampak

lima kluster yang jelas pemisahannya yaitu Kluster 1, 2 dengan 3, 4, 5 dan kluster 6. Dilihat dari sisi keragaman yang dapat dijelaskan model, pengklasteran dalam penelitian ini dapat dikategorikan baik, karena keragaman dalam kluster dapat menjelasakn sebesar 82.9% terhadap keragaman total.

Melalui Tabel 2, dapat diinterpretasikan setiap kluster yang terbentuk yaitu,

1. Kluster 1, kluster ini berkaitan dengan invensi mengenai instrumen pengukuran dan pengujian
2. Kluster 2 dan 3, berkaitan dengan invensi berhubungan dengan makanan atau bahan makanan
3. Kluster 4, kluster invensi unsur kelistrikan dasar yang berhubungan dengan pengukuran, pengujian dan teknologi mikrostruktural. Dengan kata lain alat uji/detektor mikrostruktural
4. Kluster 5, berkait dengan invensi kelistrikan dasar mengenai radar dan sejenisnya
5. Kluster 6, kluster terkait teknologi pertanian dalam arti luas
6. Kluster 7, kluster teknologi lainnya, yang merupakan kolaborasi lebih dari dua kategori teknologi.

Hasil klasterisasi setiap observasi (dokumen paten) terhadap salah satu dari ketujuh kluster dapat ditabulasi silangkan terhadap lembaga riset pemilik paten tersebut. Hal ini dilakukan untuk memetakan kecakapan tiap lembaga riset terhadap kluster teknologi yang terbentuk.

Tabel 3 memperlihatkan terjadinya indikasi adanya tumpang tindih riset antar lembaga. Kemungkinan besar terjadi tumpang tindih riset yang menghasilkan output paten, terjadi pada kluster teknologi 7. Hal ini disebabkan keempat lembaga riset tersebut sebagian paten termasuk dalam kluster teknologi 7. Temuan ini sejalan dengan penelitian terdahulu bahwa penggunaan anggaran riset diduga mengalami penyimpangan dalam bentuk penelitian fiktif, tumpang tindih aktivitas, pemotongan dana antara 10%-50%, pemberian dan penggunaannya tidak sesuai aturan, dan pengendapan anggaran (Dalilah & Pratama, 2019).

Tabel 3, dengan mengabaikan kluster 7 (kluster teknologi tumpang tindih), dapat kita identifikasi spesialisasi tiap lembaga riset berdasar kluster yang terbentuk, BATAN tampak pada Tabel 3 memiliki spesialisasi teknologi terkait radar/alat kelistrikan dasar (kluster 4). BPPT memiliki spesialisasi teknologi berhubungan dengan pertanian (Kluster 6). LAPAN secara proporsi sebagian besar teknologi terkait radar dan alat pengukuran. Sementara LIPI dari sisi proporsi, terbesar invensi terkait makanan dan bahan makanan (kluster 2 dengan 3).

Apabila diidentifikasi dari sisi banyaknya output paten terhadap kluster teknologi, terlihat bahwa LIPI merupakan lembaga riset yang paling produktif menghasilkan paten dari berbagai kluster/kategori invensi. Hal ini secara kebijakan seperti dua sisi mata uang. Pertama boleh jadi LIPI merupakan lembaga riset yang paling produktif. Kedua, pada sisi lain boleh jadi LIPI kurang memiliki spesialisasi teknologi yang jelas jika dibandingkan dengan tiga lembaga riset lainnya. Dampak strategis dari hasil temuan ini pemangku kebijakan lebih baik menciptakan ekosistem riset yang berbasis kolaborasi antar lembaga jika terdapat lebih dari satu lembaga riset pada posisi kluster teknologi yang sama. Hal ini dimaksudkan untuk menciptakan teknologi yang berdampak besar, signifikan, nyata dan tidak duplikasi. Selain itu dapat kita analisis bahwa *K-means Clustering* dengan *pre* analisis menggunakan LFSBSS dapat memberikan gambaran yang informatif dan tidak multitafsir dalam memetakan invensi lembaga riset.

Apabila dibandingkan metode *K-means Clustering* dengan *pre* analisis LFSBSS pada penelitian ini, dengan metode analisis dengan social network analisis yang dilakukan Chang (2017), terlihat bahwa metode analisis yang dilakukan dalam penelitian ini jauh lebih memudahkan dalam intepretasi dan narasi hasil analisis. Penelitian sejenis dilakukan terdahulu oleh Madani & Weber (2016), terlihat identifikasi kecakapan teknologi yang dilakukan hanya didasarkan pada banyaknya kemunculan suatu kata teknologi dalam suatu metadata paten. Metode seperti ini dirasa kurang

valid dan berisiko jika diambil sebagai dasar kebijakan, karena tanpa melihat jejaring dengan teknologi lainnya. Dengan kata lain identifikasi kecakapan teknologi sebaiknya dikaitkan dengan keterkaitan antar teknologi. Hal ini karena saat ini hampir dikatakan mustahil bahwa suatu teknologi berdiri sendiri.

Penelitian yang dilakukan oleh Lu & Liu (2016), dapat dengan baik mengidentifikasi pusat kecakapan teknologi, tetapi dalam mengintepretasikan hasil analisis cukup sulit. Hal ini dikarenakan suatu kelompok teknologi sangat rumit sekali berinteraksi dengan teknologi lainnya. Hal ini berdampak pada keragu-raguan dalam menilai dimana letak pusat kecakapan teknologi dalam suatu kluster teknologi yang terbentuk.

Penelitian terdahulu yang mengkaji terkait identifikasi pusat teknologi berbasis dokumen paten terkhusus teknologi informasi dan komunikasi pernah dilakukan Noh et al (2016). Dalam temuannya, melalui teknik analisis coupling bibliografi dan dan text mining dapat diidentifikasi pusat teknologi yang berpotensi berkembang di masa yang akan datang. Sama halnya dengan penelitian X. Wang & Duan (2011) pada dasarnya penelitian ini menerapkan keyword analisis dalam mengidentifikasi suatu pusat teknologi dalam aplikasi paten. Padahal dalam dokumen paten-sudah diidentifikasi pusat teknologi berdasarkan kode teknologi IPC WIPO yang sudah disepakati secara global.

Berdasarkan hal ini peneliti mengklaim dengan teknik analisis yang digunakan dalam penelitian ini memperbaiki kelemahan pada penelitian terkait sebelumnya. Hal ini dapat kita lihat bahwa *K-means Clustering* dengan LFSBSS, keterkaitan antar teknologi masih dapat terlihat dan aspek kemudahan intepretasi lebih sederhana dan tidak multitafsir dibandingkan *Social Network Analysis* (SNA). Penelitian ini juga menerapkan analisis pada kode IPC sehingga hasil analisis dirasa lebih valid dibandingkan dengan analisis berbasis *text mining*.

E. KESIMPULAN

Berdasarkan paparan sebelumnya dapat dilihat bahwa kluster teknologi lembaga riset

Indonesia mengkhususkan pada 1) Teknologi terkait pengembangan teknologi instrumen pengukuran dan pengujian; 2) Teknologi terkait makanan dan bahan makanan; dan 3) alat uji/detektor mikrostruktural; 4) Radar dan sejenisnya; dan 5) Teknologi dalam bidang pertanian. Hal lainnya dapat ditemukan masih terdapat indikasi tumpang tindih tema riset antar lembaga riset. Perlunya penciptaan ekosistem riset yang berbasis kolaborasi antar lembaga jika terdapat lebih dari satu lembaga riset pada posisi kluster teknologi yang sama. Hal ini ditujukan agar riset yang dihasilkan berdampak besar, signifikan, nyata, dan tidak tumpang tindih. Dilihat dampak teoritis, kajian ini menemukan bahwa peubah penjelas yang dominan belum tentu menjadi peubah kandidat penentu pemisah pembentukan kluster. Metode LFSBSS dapat menjelaskan fenomena ini dengan baik. LFSBSS mampu menyederhanakan 69 peubah penjelas menjadi hanya 8 peubah penjelas, tanpa kehilangan informasi yang berarti.

Penelitian selanjutnya disarankan untuk mengkombinasikan analisis *K-means Clustering* LFSBSS dengan metode analisis *Social Network Analysis*. Hal ini dimaksudkan untuk melihat sisi keterkaitan antar jejaring teknologi dan jejaring aktor penemu teknologi. Dengan adanya informasi jejaring teknologi dapat dilihat fenomena kolaborasi riset dan kemungkinan terciptanya disruptif teknologi.

Ucapan Terima Kasih

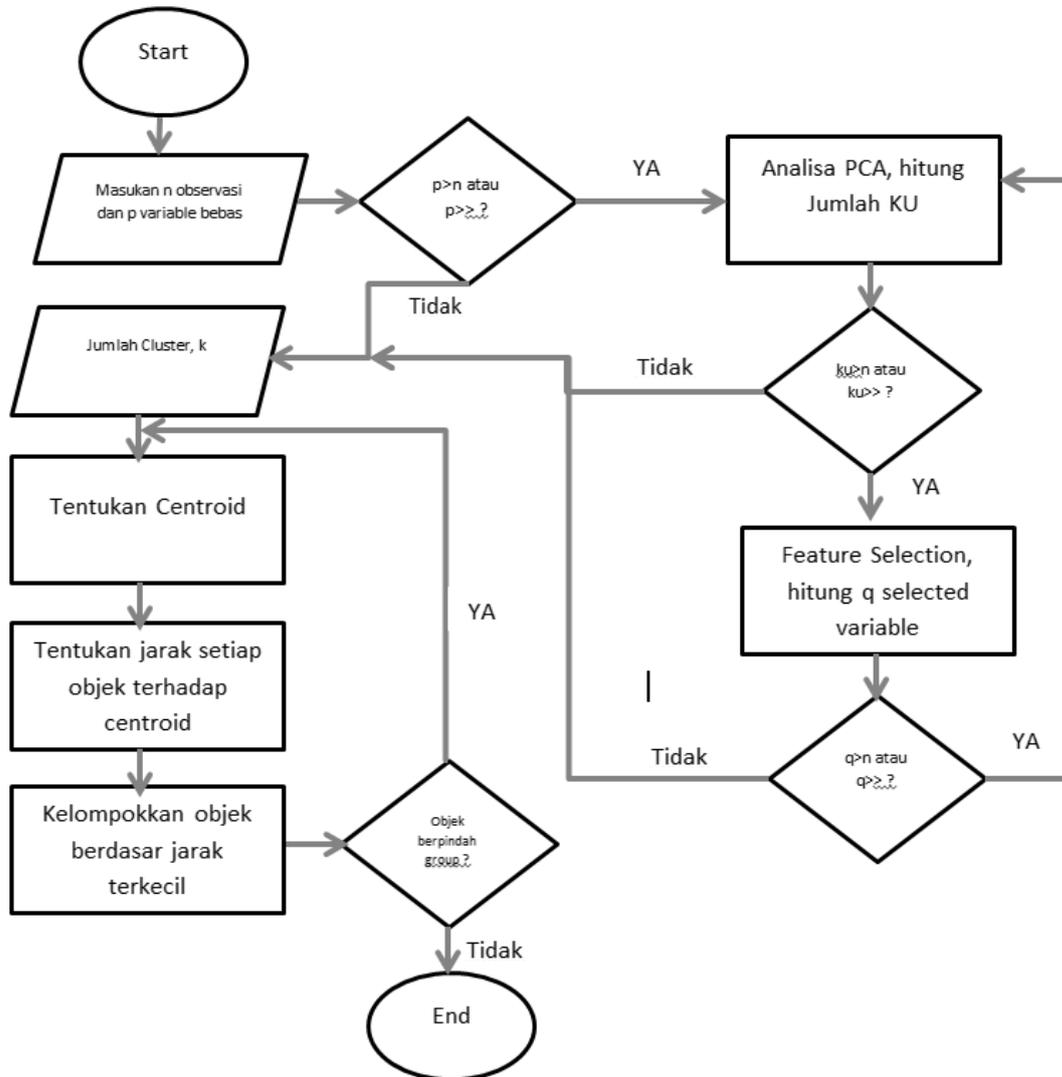
Ungkapan terima kasih disampaikan kepada Kementerian Riset dan Teknologi/BRIN atas program Beasiswa SAINTEK, kesempatan dan fasilitas yang diberikan.

DAFTAR PUSTAKA

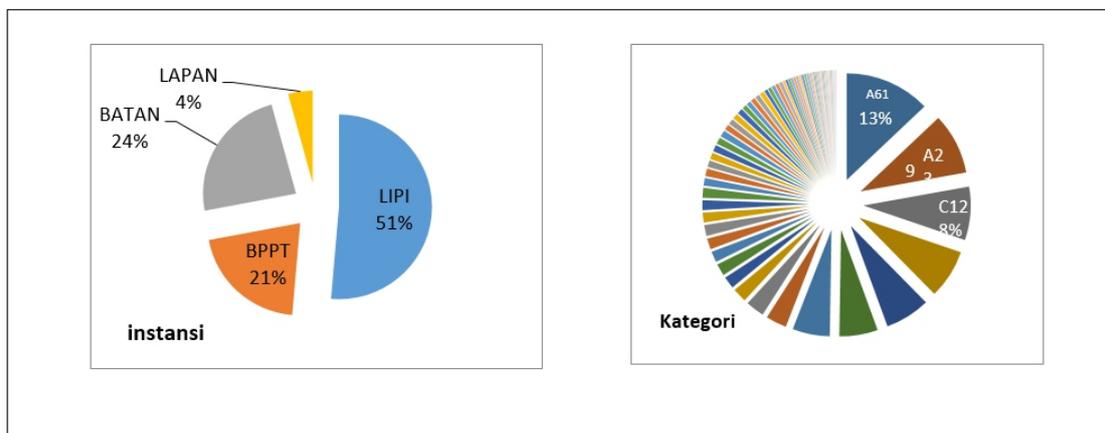
- Aggarwal, Charu C.; Reddy, C. K. (2014). *Data clustering algorithms and applications (Whole Book)* (C. K. Aggarwal, Charu C.; Reddy (ed.)). Taylor & Francis.
- Artz, K. W., Norman, P. M., Hatfield, D. E., & Cardinal, L. B. (2010). A longitudinal study of the impact of R&D, patents, and product innovation on firm performance. *Journal of Product Innovation Management*, 27(5), 725–740. <https://doi.org/10.1111/j.1540-5885.2010.00747.x>
- Chang, S. (2017). Technological forecasting & social change the technology networks and development trends of university-industry collaborative patents. *Technological Forecasting & Social Change*, 118, 107–113. <https://doi.org/10.1016/j.techfore.2017.02.006>
- Chen, H., Zhang, G., Zhu, D., & Lu, J. (2017). Topic-based technological forecasting based on patent data: A case study of Australian patents from 2000 to 2014. *Technological Forecasting and Social Change*, 119, 39–52. <https://doi.org/10.1016/j.techfore.2017.03.009>
- Dalilah, E., & Pratama, F. (2019). Permasalahan dan rumusan perbaikan pengelolaan dana penelitian di Indonesia. *Integritas*, 6(1), 109–124.
- Dhir, S., Kumar, D., & Singh, V. B. (2019). Success and failure factors that impact on project implementation using agile software development methodology. In M. . Hoda (Ed.), *Advances in Intelligent Systems and Computing* (731st ed., Vol. 731). Springer. https://doi.org/10.1007/978-981-10-8848-3_62
- Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistics* (6th ed.). Pearson Prentice Hall.
- Kim, G., & Bae, J. (2017). A novel approach to forecast promising technology through patent analysis. *Technological Forecasting and Social Change*, 117, 228–237. <https://doi.org/10.1016/j.techfore.2016.11.023>
- Li, Y., Dong, M., & Hua, J. (2008). Localized feature selection for clustering. *Pattern Recognition Letters*, 29(1), 10–18. <https://doi.org/10.1016/j.patrec.2007.08.012>
- Lu, L. Y. Y., & Liu, J. S. (2016). Technological Forecasting & social change a novel approach to identify the major research themes and development trajectory : The case of patenting research. *Technological Forecasting & Social Change*, 103, 71–82. <https://doi.org/10.1016/j.techfore.2015.10.018>

- Madani, F., & Weber, C. (2016). The evolution of patent mining : Applying bibliometrics analysis and keyword network analysis. *World Patent Information*, 46, 32–48. <https://doi.org/10.1016/j.wpi.2016.05.008>
- Markowitz, J. S. (2018). Multivariate analysis. *Statistics Reference Online*, 71–81. https://doi.org/10.1007/978-3-319-77203-5_8
- Mubarok, M. F. (2017). *Top 10 Perusahaan Farmasi Terbesar Indonesia*. <https://farmasiindustri.com>. <https://farmasiindustri.com/industri/top-10-perusahaan-farmasi-indonesia.html>
- Noh, H., Song, Y. K., & Lee, S. (2016). Identifying emerging core technologies for the future: Case study of patents published by leading telecommunication organizations. *Telecommunications Policy*, 40(10–11), 956–970. <https://doi.org/10.1016/j.telpol.2016.04.003>
- Presiden Republik Indonesia. (2016). *Undang-Undang No 13 Tahun 2016. Paten* (Issue 1). http://www.dgip.go.id/images/ki-images/pdf-files/uu_pp1/UU-nomor-13-tahun-2016-tentang-paten.pdf
- Presiden RI. (2019). *Peraturan Presiden Republik Indonesia No 74 Tahun 2019 tentang Badan Riset dan Inovasi Nasional* (Issue 009525).
- Schwab, K. (2019). *The Global Competitiveness Report 2019*. http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf
- Su, F. P., Chen, S. J., Chang, Y. H., & Lai, K. K. (2019). Construct a three-stage analysis model of integrated main path analysis and patent family-exploring the development of blockchain. *ACM International Conference Proceeding Series*, 151–156. <https://doi.org/10.1145/3374549.3374583>
- Wang, B., Liu, Y., Zhou, Y., & Wen, Z. (2018). Emerging nanogenerator technology in China: A review and forecast using integrating bibliometrics, patent analysis and technology roadmapping methods. *Nano Energy*, 46, 322–330. <https://doi.org/10.1016/j.nanoen.2018.02.020>
- Wang, X., & Duan, Y. (2011). Identifying core technology structure of electric vehicle industry through patent co-citation information. *Energy Procedia*, 5, 2581–2585. <https://doi.org/10.1016/j.egypro.2011.03.443>
- Wang, Y. L. (2012). Research on technology selection for enterprises with tools of patent analysis. *International Conference on Management Science and Engineering - Annual Conference Proceedings*, 1, 1651–1657. <https://doi.org/10.1109/ICMSE.2012.6414394>
- WIPO. (2018). Guide to the international patent classification. *WIPO (World Intellectual Property Organization)*. https://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc.pdf
- Yu, X., & Zhang, B. (2019). Obtaining advantages from technology revolution: A patent roadmap for competition analysis and strategy planning. *Technological Forecasting and Social Change*, 145(April), 273–283. <https://doi.org/10.1016/j.techfore.2017.10.008>
- Zelterman, D. (2015). Applied multivariate statistics with R. In M. Gail, J. M. Samet, A. Tsiatis, & W. Wong (Eds.), *Applied Multivariate Statistics with R*. springer. <https://doi.org/10.1007/978-3-319-14093-3>

DAFTAR GAMBAR

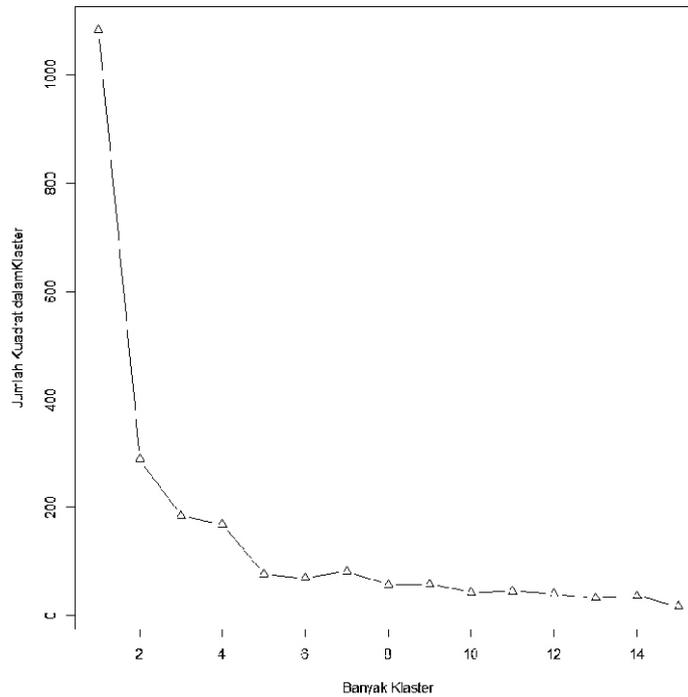


Gambar 1. Diagram Alir Analisis



Gambar 2. Distribusi Paten Nasional

DAFTAR GAMBAR



Gambar 3. Scree Plot Penentuan Jumlah Kluster (Sumber : Data diolah, 2019)

DAFTAR TABEL

Tabel 1. Localized feature selection for clustering

Peubah proposed	Type of step	BICclust	Model	G	BICdiff	Decision
H01	Add	737.7576	E	5	1208.596	Accepted
A23	Add	1751.3041	EII	9	1584.0388	Accepted
A01	Add	2939.7359	EII	13	1580.4901	Accepted
A01	Remove	1751.3041	EII	9	1580.4901	Rejected
C08	Add	4066.1556	EII	20	1611.3606	Accepted
C08	Remove	2939.7359	EII	13	1611.3606	Rejected
B27	Add	4656.4015	EII	20	205.5406	Accepted
B27	Remove	4066.1556	EII	20	205.5406	Rejected
C11	Add	5392.9184	EII	20	351.8116	Accepted
C08	Remove	5985.0568	EII	15	-107.1974	Accepted
G01	Add	7250.6849	EII	19	1674.2668	Accepted
B27	Remove	5669.7642	EII	18	1196.2153	Rejected
G10	Add	8867.2323	EII	20	1231.8421	Accepted
B27	Remove	7250.6849	EII	19	1231.8421	Rejected
B81	Add	10575.7184	EII	20	935.6184	Accepted
B81	Remove	8867.2323	EII	20	935.6184	Rejected
G06	Add	10537.1101	EII	20	-426.3406	Rejected
B81	Remove	8867.2323	EII	20	935.6184	Rejected

Sumber: Data primer diolah, 2019

DAFTAR TABEL

Tabel 2. Rangkuman Analisis Kluster

Klaster Means	H01	A23	A01	B27	C11	G01	G10	B81
1	0.00	0.00	0.00	0.00	0.00	2.08	0.00	0.00
2	0.00	1.43	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	3.29	0.00	0.00	0.00	0.00	0.00	0.00
4	1.33	0.00	0.00	0.00	0.00	0.07	0.00	0.07
5	3.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	2.30	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.05	0.01	0.01	0.08	0.01	0.00

Within klaster sum of squares by klaster:

[1] 0.9230769 5.6521739 3.4285714 5.2000000 2.7500000 4.1000000 37.9711538

(between_SS / total_SS = 82.9 %)

Sumber: Data primer diolah, 2019

Tabel 3. Tabulasi Silang Kluster Teknologi terhadap Pemilik Paten

Klaster	Jumlah Paten			
	BATAN	BPPT	LAPAN	LIPI
1	4	2	2	5
2	0	3	0	20
3	0	2	0	5
4	6	2	2	5
5	1	0	0	3
6	2	4	0	4
7	49	41	5	113

Sumber : Data primer diolah, 2019