

Analisis Cluster Data Interkomparasi Anak Timbangan dengan Algoritma *Self Organizing Maps*

<http://dx.doi.org/10.28932/jutisi.v7i2.3698>

Riwayat Artikel

Received: 18 Juni 2021 | Final Revision: 15 Juli 2021 | Accepted: 23 Juli 2021

Arif Fajar Solikin[✉]#1, Kusrini^{*}2, Ferry Wahyu Wibowo^{#3}

[#] Magister Teknik Informatika, Universitas AMIKOM Yogyakarta

Jl. Ring Road Utara, Ngringin, Condongcatu, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281

¹arif.1311@students.amikom.ac.id

³Ferry.w@amikom.ac.id

^{*} Magister Teknik Informatika, Universitas AMIKOM Yogyakarta

Jl. Ring Road Utara, Ngringin, Condongcatu, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281

²kusrini@amikom.ac.id

Abstract — Intercomparison was conducted to determine the ability and the performance of the laboratory. Intercomparison results are usually expressed in the range of En ratio values ($En \leq 1$) which express the equivalence of one laboratory with other laboratories. If the laboratory is declared unequal, then it needs to identify the source of the problem by itself. To make it easier, it can be done by Clustering which is one of the data mining techniques. Clustering is done by applying a self organizing map algorithm on the KNIME (Konstanz Information Miner) analytic tools. Several experiments were carried out with different layer size and data normalization status from one experiment to another experiment. The results were analyzed through pseudo F statistical test and icdrate test. The largest pseudo F statistic value was obtained from the 8th experiment (setting the layer size 2x2 without data normalization) with a pseudo F statistic value of 167.53 for 1kg artifacts and a Pseudo F statistic value of 104.86 for 200 g artifacts where the optimum number of clusters are 4. The smallest icdrate value was obtained from the 5th experiment (setting the 2x3 layer size without data normalization) with an icdrate value of 0.0713 for 1kg artifacts and icdrate value of 0.2889 for 200g artifacts with the best number of clusters being 6. From 12 laboratories can be grouped into 6 groups where each group has the same identification. There are groups 1, 3 and 6 have 1 member, while groups 2, 4 and 5 have 3 members.

Keywords— Cluster; Intercomparison; Optimum.

I. PENDAHULUAN

Transaksi perdagangan yang dilakukan antara pembeli dan penjual biasanya melibatkan alat ukur dalam penentuan kuantitas barang / jasa. Nilai rupiah dari barang / jasa yang diserahkan dari penjual ke pembeli dipengaruhi oleh nilai berat atau kuantitas barang yang ditentukan dari alat ukur milik penjual. Alat ukur tersebut disahkan oleh Unit Metrologi Legal Pemerintah Kabupaten / Kota berdasarkan Undang-Undang No 23 Tahun 2014 tentang Pemerintah Daerah. Balai Standardisasi Metrologi Legal (BSML) sebagai Unit Pelaksana Teknis Kemetrolgian di Kementerian Perdagangan memiliki tugas pokok salah satunya berupa verifikasi standar dan uji banding laboratorium metrologi legal (Peraturan Menteri Perdagangan No 81 Tahun 2020). Terdapat 4 (empat) BSML yang terdiri dari BSML Regional I sampai dengan BSML Regional IV. Di wilayah regional 2 yang meliputi Jawa, Bali, Nusa Tenggara Barat dan Nusa Tenggara Timur, terdapat 125 Unit Metrologi Legal yang sudah berdiri. Dari jumlah 125 kantor unit metrologi legal tersebut, terdapat 36 kantor unit metrologi legal yang telah melakukan pengelolaan laboratorium metrologi legal. Untuk mengetahui unjuk kerja serta kemampuan unit metrologi legal dalam melaksanakan verifikasi standar secara mandiri, dilakukanlah interkomparasi antar laboratorium metrologi legal. Interkomparasi dilakukan pada besaran massa yaitu anak timbangan dimana hasil interkomparasi berupa data yang memuat massa konvensional uji anak timbangan beserta nilai ketidakpastian pada level kepercayaan 95%. Data ini kemudian dilakukan perhitungan dengan menggunakan metode derajat kesetaraan (*Degrees of Equivalent / DoE*) atau yang lebih dikenal sebagai rasio En yang mengacu pada ISO/IEC 17043:2010[1]. Hasilnya berupa pernyataan kesetaraan laboratorium peserta interkomparasi, laboratorium dinyatakan setara jika nilai $En \leq 1$. Selama ini

laboratorium akan menerima hasil berupa nilai En dengan status inlier atau outlier. Untuk status inlier umumnya laboratorium dalam kondisi normal seperti laboratorium acuan, namun jika statusnya outlier laboratorium perlu melakukan investigasi sumber masalahnya secara mandiri.

Untuk mendapatkan analisis hasil interkomparasi secara mendalam, dirancang satu penelitian dengan melakukan *Clustering* data hasil interkomparasi menggunakan algoritma *self organizing maps*. *Clustering* atau pengelompokan data dilakukan supaya data hasil interkomparasi dari masing-masing laboratorium yang memiliki permasalahan yang sama atau sumber ketidakpastian yang sama dapat diidentifikasi. Identifikasi perlu dilakukan oleh penyelenggara interkomparasi agar dapat memberikan bimbingan dan rekomendasi perbaikan pengelolaan laboratorium metrologi legal kepada unit metrologi legal yang menjadi peserta interkomparasi. Penggunaan algoritma *self organizing maps* didasarkan pada cara kerja algoritma dalam menyederhanakan dimensi data dan hubungan antar *neuron* pada layer *input* dan *output* yang secara terus menerus memperbarui bobotnya sehingga diperoleh kesamaan ciri antar anggota dalam *cluster* yang sama [2].

Pemilihan algoritma *self organizing maps* didasarkan pada penelitian yang telah dilakukan sebelumnya. Penelitian [3] penerapan algoritma *self organizing maps* dan *k* means dengan $k=2$ dimana bobot awal yang ditentukan adalah sama untuk kedua metode telah menghasilkan ketepatan klasifikasi yang sama. Proses *Clustering* dilakukan dengan menentukan *vector node* yang dipilih untuk selanjutnya ditentukan BMU (*best matching unit*) nya dengan *node* lain. Penelitian [4] diawali dengan data yang dinormalisasi pada rentang 0 sampai dengan 1. Data yang sudah dinormalisasi kemudian dilakukan *Clustering* dengan menetapkan jumlah iterasi dan nilai *learning rate*-nya. Penelitian lain [5] terkait pembuatan aplikasi *Clustering* data IPM (indeks pembangunan manusia) Aceh menggunakan *bubble function* sebagai *neighborhood function* dimana radius *neighborhood* sama dengan nol. Fungsi *bubble* dipilih dengan alasan proses komputasi lebih cepat dan kuantisasi kesalahan lebih kecil. Penelitian selanjutnya [6] tentang pengelompokan wilayah berdasarkan kesejahteraan sosial menggunakan *Silhouette Coefficient* untuk mengukur kualitas hasil *Clustering* dengan membatasi nilai rentang lebih dari 1. Penelitian [7] yang dilakukan oleh Sivakkumaran Lakshminarayanan, mencoba menentukan pengelompokan penggunaan BBM berdasar data manuver pengemudi kendaraan dengan menggunakan perluasan algoritma SOM yaitu *super organizing maps*, dimana data *input* dibedakan menjadi 4 (empat) layer, masing-masing layer kemudian dilakukan proses pengelompokan. Penelitian terakhir [8] oleh Jin Tian dan Mengyi Gu menentukan subruang pada *neuron* dengan tujuan mereduksi *noise* pada data. Algoritma SOM menghasilkan *cluster* yang mungkin dari setiap dimensi, kemudian dilakukan proses penggabungan untuk menggabungkan file kelompok *neuron* dan subruang yang sesuai dengan dimensi terkait. Hasil pengelompokan akhir disimpulkan dari *cluster neuron* yang terbentuk dengan mengganti setiap *neuron* dengan poin data di dalamnya. Penelitian-penelitian tersebut telah menerapkan algoritma *self organizing maps* untuk pengelompokan data dengan tujuan yang berbeda-beda. Pada penelitian ini dilakukan *Clustering* data interkomparasi, data interkomparasi satu laboratorium dengan laboratorium yang lainnya dinyatakan memiliki kesetaraan nilai berdasarkan evaluasi nilai rasio En. Kelompok yang dihasilkan dari proses *Clustering* ditentukan oleh berapa banyak anggota *cluster* yang memiliki kedekatan atau kesamaan ciri dengan anggota lainnya. Dari hasil pengelompokan, selanjutnya akan dilakukan analisis mendalam dengan melakukan *clustering* pada sumber ketidakpastian (*repeatability*, *drift* standar dan standar yang digunakan, koreksi udara/*bouyancy*, dan resolusi timbangan).

Clustering merupakan metode yang digunakan dalam pengelompokan data. Data dikelompokkan ke beberapa *cluster* sesuai dengan kesamaan ciri atau homogenitasnya. Hasil *Clustering* yang baik akan menghasilkan tingkat kesamaan yang tinggi dalam satu *cluster* dan tingkat kesamaan yang rendah diantara *cluster* [2]. Kesamaan tersebut berupa hasil pengukuran secara numerik terhadap dua buah objek. Objek yang memiliki kemiripan tinggi maka akan memiliki kesamaan yang tinggi. Begitu sebaliknya, objek dengan kemiripan rendah akan berbeda penempatan dalam *clusternya*. Kualitas hasil *Clustering* sangat tergantung pada metode yang dipakai.

Algoritma *Clustering* dinyatakan baik jika memenuhi hal-hal berikut ini [2]:

1. Skalabilitas, dimana metode klastering mampu diperluas untuk sekian banyak data,
2. Mampu menganalisa semua jenis data,
3. Hasil *cluster* yang terbentuk bisa berbeda beda,
4. Mampu mengatasi *noise* data
5. Sensitivitas terhadap perubahan input, dimana perubahan data yang rendah akan mengakibatkan perubahan yang mencolok,
6. Dapat mengklaster data berdimensi tinggi
7. Dapat diinterpretasi dan berguna

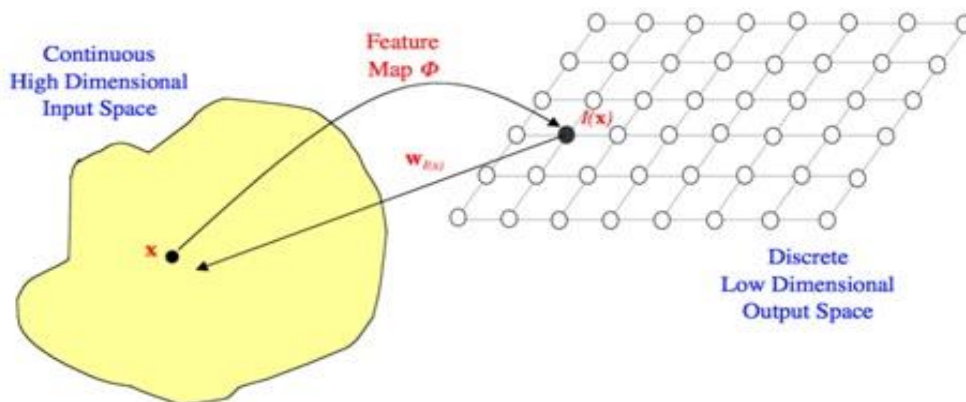
Secara umum terdapat dua pendekatan yang digunakan dalam klastering, yaitu pendekatan hirarki dan pendekatan partisi. Pendekatan hirarki dilakukan dengan membuat dendogram, data yang memiliki kesamaan ciri dikelompokkan pada hirarki yang tidak berjauhan, sedangkan yang tidak ada kesamaan ciri ditempatkan pada hirarki yang tidak berdekatan. Sedangkan pada pendekatan partisi, data dikelompokkan dengan cara memilah milah data ke dalam klaster yang ada secara optimal tanpa melihat hirarki satu dan lainnya. Pada metode partisi setiap klaster memiliki titik pusat klaster (*centroid*) dimana *centroid* ini memiliki jarak terkecil dibanding *neuron* lainnya.

Self-organizing maps (SOM) awal mulanya dikenalkan oleh Profesor Teuvo Kohonen pada tahun 1996. SOM merupakan teknik dalam *neural network* yang menggambarkan data dengan cara mengurangi dimensi data, data dengan dimensi tinggi dipetakan ke data berdimensi rendah agar mudah dipahami. Algoritma ini dikenal juga sebagai algoritma tidak terawasi atau tidak terbimbing. Tidak memerlukan data training dalam proses pengelompokan datanya.

Pada algoritma SOM, setiap *neuron input* yang diberikan inisial bobot w akan saling menyesuaikan nilai jaraknya sampai diperoleh nilai minimum pada fungsi diskriminan. Selama proses penyesuaian diri atau dikenal dengan istilah *self-organizing*, *neuron* dengan jarak terkecil, dimana memiliki fungsi diskriminan paling minim akan dipilih sebagai pemenang. *Neuron* pemenang dan *neuron* sekitar akan memperbarui bobot mereka secara terus menerus. Setiap output akan bereaksi terhadap pola input tertentu yang pada akhirnya menghasilkan kesamaan karakteristik antar anggota dalam *cluster* yang sama [9].

Dalam jaringan SOM, *neuron target* diletakkan tidak dalam satu baris seperti pada model JST lainnya. *Neuron target* terletak pada bentuk dua dimensi, di mana bentuknya dapat di set sesuai kebutuhan. Struktur topologi yang berbeda akan menghasilkan *neuron* di sekitar *neuron* pemenang yang berbeda, sehingga bobotnya juga akan berbeda. Dalam SOM, perubahan bobot tidak hanya terjadi pada jarak ke *neuron* pemenang, tetapi juga jarak dengan *neuron* sekitarnya. *Neuron* yang mengelilingi *neuron* pemenang dipilih berdasar jaraknya dari *neuron* pemenang.

Secara umum, cara kerja SOM ditunjukkan oleh Gambar 1 berikut:



Gambar 1 Cara kerja *Self Organizing Maps* [2].

Titik (x) di ruang berdimensi tinggi diproyeksikan ke titik $I(x)$ di ruang berdimensi rendah. Setiap titik (I) di ruang output akan dipetakan ke titik yang sesuai di ruang input dengan bobot $w_{I(x)}$.

Terdapat tiga komponen penting pada algoritma SOM [2], yaitu:

1. Persaingan: Pada setiap fungsi masukan, *neuron* akan menyesuaikan jaraknya dengan *neuron* lain setelah diberikan bobot tertentu, terjadi persaingan antar *neuron*. Persaingan tersebut akan menghasilkan *neuron* pemenang.
2. Kerjasama: *Neuron* pemenang menentukan tempat pada lingkungan topologinya, terjadi kerjasama antar *neuron* dalam lingkungan yang sama.
3. Adaptasi sinaptik: *Neuron* tereksitasi menyesuaikan dengan bobot yang ada untuk mengurangi nilai fungsi diskriminan, sehingga *neuron* yang menang akan meningkatkan respons ke aplikasi berikutnya dengan fungsi masukan yang sama.

Pengelompokan data menggunakan algoritma SOM terdiri dari 4 tahap [2], yaitu:

1. Kompetisi: *neuron* output j , dihitung nilai jarak $D(x, w_i)$ dimana x dan w_i adalah jarak *euclydean*. Jarak *euclydean* dirumuskan dengan persamaan 1 berikut ini:

$$D(x, w_i) = \sqrt{\sum_{i=1}^n (x_i - w_{mi})^2} \quad (1)$$

dimana x adalah vector dari node input, sedangkan w_m adalah vector bobot dari node *neuron* ke m

2. Update bobot: setelah diperoleh nilai jarak dari masing-masing *neuron* input ke vector bobot, nilai jarak terkecil ditetapkan sebagai *neuron* pemenang. Setiap *neuron* pemenang beserta *neuron* sekitarnya melakukan proses penyesuaian dengan cara memperbaharui nilai bobot dimana $h(t)$ adalah *neighborhood function* dan t adalah jumlah perulangan. Fungsi *node* tetangga merupakan fungsi *Gaussian* yang direpresentasikan dengan persamaan 2 berikut ini :

$$h(t) = \alpha(t)x e^{\left(\frac{\|ri-rc\|^2}{2\delta^2(t)}\right)} \quad (2)$$

dimana

$\alpha(t)$ adalah nilai learning rate atau dikenal juga dengan nilai alpha. Nilai *learning rate* merupakan fungsi penurunan tingkat pembelajaran seiring perubahan waktu. Nilai *learning rate* diperoleh dari:

$$\alpha(t) = \alpha_i \left(1 - \frac{t}{t_{max}}\right) \quad (3)$$

α_i adalah nilai awal ke i dan

t_{max} adalah jumlah perulangan maksimum.

$\|ri - rc\|^2$ adalah jarak kuadrat antara *neuron* ke- i dengan *neuron* pemenang pada lingkungan topologi tertentu. Neuron atau data input pada clustering SOM masing-masing akan menghitung jaraknya dengan neuron pemenang (r_c) pada masing-masing grid atau layer. Masing - masing data akan menyesuaikan diri dan berkelompok dengan neuron pemenang dalam satu grid.

$\delta^2(t)$ adalah lebar tetangga yang diperoleh dari perhitungan berikut ini:

$$\delta(t) = \delta_i \left(\frac{\delta_f}{\delta_i}\right)^{\frac{t}{t_{max}}} \quad (4)$$

$\delta(t)$ adalah lebar tetangga yang seiring waktu akan berkurang jumlah perulangannya.

δ_i adalah nilai awal lebar tetangga dan

δ_f nilai akhir lebar tetangga.

3. Selalu memperbaiki nilai α dan δ . α dan δ dilakukan perubahan nilai *learning rate* dan nilai lebar tetangga sesuai persamaan 3 dan 4.
4. Menghentikan percobaan: percobaan dihentikan jika sejumlah kriteria telah tercapai. Kriteria tersebut ditentukan berdasarkan banyaknya perulangan, nilai *minimum error*, nilai alpha dan lebar tetangga.

Interkomparasi atau *interlaboratory comparison* adalah organisasi, kinerja dan evaluasi pengukuran atau tes pada dua laboratorium atau lebih sesuai dengan kondisi yang telah ditentukan [1]. Interkomparasi atau disebut juga dengan istilah uji profisiensi telah menjadi aspek penting dalam praktik laboratorium di semua bidang pengujian, kalibrasi, dan inspeksi. Sebagian besar skema uji profisiensi memiliki ciri umum berupa perbandingan hasil dari satu laboratorium dengan hasil dari satu atau lebih laboratorium yang berbeda. Sifat pengujian atau pengukuran yang dilakukan dalam interkomparasi dimaksudkan untuk membandingkan kinerja laboratorium. Dapat disimpulkan bahwa tujuan dari interkomparasi adalah untuk mengetahui kemampuan atau unjuk kerja dari suatu laboratorium.

Ada tiga jenis pemeriksaan laboratorium dasar: kuantitatif, kualitatif dan interpretatif [14].

1. Hasil pengukuran kuantitatif adalah numerik dan dilaporkan dalam skala interval atau rasio. Pengujian untuk pengukuran kuantitatif dapat bervariasi dalam presisi, kebenaran, kepekaan analitis, dan kekhususan. Dalam skema pengujian profisiensi kuantitatif, hasil numerik biasanya dianalisis secara statistik.
2. Hasil tes kualitatif bersifat deskriptif dan dilaporkan dalam skala kategorikal atau ordinal, misalnya identitas mikro-organisme, atau dengan identifikasi adanya ukuran tertentu (seperti obat atau penilaian suatu karakteristik). Penilaian kinerja dengan analisis statistik mungkin tidak sesuai untuk pemeriksaan kualitatif.
3. Dalam tes interpretatif, uji profisiensi berupa hasil tes seperti pernyataan morfologi deskriptif, atau satu set data (misalnya untuk menentukan garis kalibrasi) atau satu set informasi lainnya misalnya studi kasus tentang fitur interpretatif dari kompetensi peserta.

Hasil pengukuran numerik dari laboratorium diolah secara statistik dengan menggunakan nilai rasio E_n . Nilai rasio E_n dihitung dengan rumus:

$$E_n = \left| \frac{x_i - x_{ref}}{\sqrt{U_i^2 + \sqrt{U_{ref}^2}}} \right| \quad (5)$$

Dimana,

x_i merupakan nilai koreksi anak timbangan pada laboratorium peserta dengan i merujuk pada urutan no peserta interkomparasi.

x_{ref} merupakan nilai koreksi anak timbangan pada laboratorium referensi.

U_i merupakan ketidakpastian laboratorium peserta pada level kepercayaan 95%.

U_{ref} merupakan ketidakpastian laboratorium referensi pada level kepercayaan 95%.

Hasil perhitungan nilai En berupa pernyataan kesetaraan laboratorium peserta interkomparasi. Laboratorium dinyatakan setara jika nilai $En \leq 1$.

Untuk mengetahui jumlah kelompok optimum dapat digunakan kriteria nilai *Pseudo F-statistics* [10]. Berikut adalah persamaan yang digunakan untuk menghitung nilai *Pseudo F-statistics*:

$$Pseudo F Statistic = \frac{\left(\frac{R^2}{k-1}\right)}{\left(\frac{1-R^2}{n-k}\right)} \quad (6)$$

Dimana:

$$R^2 = \frac{(SST-SSW)}{SST} \quad (7)$$

$$SST = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (X_{ijk} - \underline{X}_j)^2 \quad (8)$$

$$SSW = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (X_{ijk} - \underline{X}_{jk})^2 \quad (9)$$

Keterangan:

SST (Sum Square Total): Total jumlah dari kuadrat jarak sampel terhadap rata-rata keseluruhan

SSW (Sum Square Within): Total jumlah dari kuadrat jarak sampel terhadap rata-rata kelompoknya

n : banyaknya sampel

c : banyaknya variabel

p : banyaknya kelompok

x_{ijk} : sampel ke- i pada variabel ke- j kelompok ke- k

x_j : rata-rata seluruh sampel pada variabel ke- j

x_{jk} : rata-rata sampel pada variabel ke- j dan kelompok ke- k

Nilai *pseudo F-statistics* tertinggi menunjukkan bahwa jumlah kelompok yang digunakan untuk mempartisi data telah optimal. Dimana keragaman dalam kelompok sangat homogen sedangkan antar kelompok sangat heterogen [10].

Klaster terbaik untuk mengelompokkan data memiliki nilai homogenitas antar anggota dalam satu klaster yang tinggi dan memiliki nilai heterogenitas antar klaster yang tinggi. Dengan kata lain variasi antar anggota dalam satu klaster harus rendah sedangkan variasi antar klaster harus tinggi. Untuk menentukan kriteria klaster terbaik dapat digunakan nilai *icdrate*. *Icdrate (internal cluster dispersion)* menggambarkan tingkat dispersi atau perbedaaan dalam klaster. Nilai *icdrate* yang semakin kecil menunjukkan bahwa klaster tersebut semakin baik sebab antara anggota dalam satu klaster memiliki perbedaan yang rendah atau memiliki variasi yang kecil [10].

Nilai *icdrate* dapat diperoleh dengan menggunakan persamaan:

$$icdrate = 1 - \frac{SSB}{SST} \quad (10)$$

Dimana:

$$SSB = \sum_{j=1}^c \sum_{k=1}^p (\underline{X}_{jk} - \underline{X}_j)^2 \quad (11)$$

Keterangan:

SSB (*Sum Square Between*): Total jumlah dari kuadrat jarak rata-rata sampel pada variabel ke- j dan kelompok ke- k terhadap rata-rata sampel pada variabel ke- j

\bar{X}_{jk} : rata-rata sampel pada variabel ke- j dan kelompok ke- k

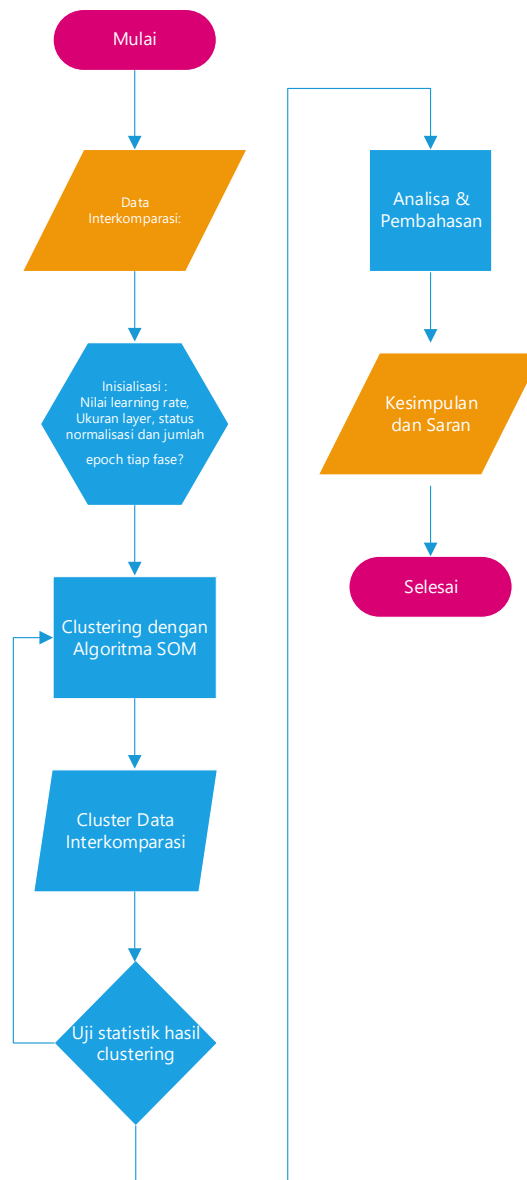
\bar{X}_j : rata-rata sampel pada variabel ke- j

c : banyaknya variabel

p : banyaknya kelompok

II. METODE PENELITIAN

Penelitian ini merupakan penelitian eksperimen dengan pendekatan kuantitatif, dimana data penelitian berupa data numerik hasil dari pengujian artefak interkomparasi yang dianalisis dengan metode statistik. Data interkomparasi anak timbangan terlebih dahulu dihitung nilai rasio En nya sebelum dilakukan *Clustering* dengan algoritma *self organizing maps*. Adapun tahapan – tahapan penelitian yang akan dilakukan dijelaskan dalam gambar 2 sebagai berikut:



Gambar 2 Alur Penelitian

Pada gambar 2 di atas, setelah dilakukan *literature review* untuk mengumpulkan bahan referensi dan tinjauan pustaka yang bersumber dari jurnal internasional dan jurnal nasional, maka selanjutnya dilakukan tahapan sebagai berikut:

1. Tahap pengumpulan data interkomparasi anak timbangan. Data ini diperoleh dari kegiatan interkomparasi tahun 2021 yang diselenggarakan oleh Balai Standardisasi Metrologi Regional II dengan jumlah peserta sebanyak 36 laboratorium[15]. Data yang dikumpulkan baru data loop I yang terdiri dari 12 laboratorium (loop II dan loop III mulai dilaksanakan Juli 2021) berupa nilai massa konvensional anak timbangan dan nilai *budget uncertainty* untuk artefak 1 kg dan artefak 200 g seperti tampak pada tabel 1 dan tabel 2 . Dari data yang terkumpul kemudian dilakukan perhitungan nilai rasio En nya. Data Nilai rasio En dan data *budget uncertainty* dari laboratorium peserta interkomparasi akan dilakukan proses pengelompokkan data dengan menggunakan algoritma SOM.
2. Sebelum *clustering* dilakukan terlebih dahulu ditentukan nilai parameter dari nilai *learning rate*, Jumlah *epoch* pada fase *order* dan fase *convergen*, dan nilai tinggi lebarnya *lattice / layer*.
3. Tahap pengelompokkan data. Setelah inisialisasi, selanjutnya dilakukan proses *clustering* dengan algoritma SOM. Proses *clustering* dilakukan dengan menggunakan *platform KNIME analytic*. Hasil *clustering* kemudian dievaluasi secara statistik dengan menghitung nilai *pseudo F statistic* dan nilai *internal cluster dispersion (icdrate)*. Nilai yang dicari adalah nilai *pseudo F statistic* terbesar dan nilai *icdrate* yang mendekati nol tapi tidak kurang sama dengan nol ($>=0$). Jika nilai yang dicari masih belum sesuai baik *pseudo F statistic* ataupun *icdrate*, maka dilakukan inisialisasi ulang nilai parameter dan dilakukan *clustering* kembali.
4. Setelah didapatkan hasil Clustering kemudian dilakukan Analisis hasil Clustering. Analisis dilakukan dengan menghitung nilai Pseudo F Statistic dan nilai *icdrate*.
5. Tahapan terakhir dalam penelitian ini adalah menarik kesimpulan yang diperoleh dari eksperimen dan memberikan saran untuk penelitian selanjutnya.

TABEL 1
DATA INTERKOMPARASI ARTEFAK 1 KG

Kode	Mc 1kg	Uc 1kg	En	Ur	Ud	Ustd	Udrif	Ub
A1 (acuan)	1000.0053	0.0031		0.0000416	0.0000408	0.0008	0.000739	0.0011085
A2	999.99	0.0087	1.2966	0.000000	0.004083	0.000800	0.000739	0.001109
A3	1000.0170	0.012	0.7748	0.004082	0.002300	0.010607	0.000000	0.001109
A4	1000.0025	0.011	0.1947	0.002850	0.004080	0.002300	0.000740	0.001110
A5	1000.0039	0.0090	0.1157	0.001300	0.004100	0.000800	0.000200	0.001100
A6	1000.0394	0.0030	5.5902	0.000200	0.000400	0.000600	0.000700	0.001100
A7	1000.0013	0.0098	0.3101	0.000000	0.004080	0.001480	0.002310	0.000069
B1 (acuan)	1000.0037	0.0031		0.000042	0.000041	0.000800	0.000739	0.001109
B2	1000.0006	0.0087	0.2585	0.000000	0.004100	0.000800	0.000800	0.001100
B3	1000.0036	0.0091	0.0041	0.001900	0.004100	0.000800	0.000200	0.069300
B4	1000	0.011	0.2589	0.003350	0.004082	0.000145	0.000739	0.001109
B5	999.996	0.010	0.5840	0.002500	0.004100	0.000600	0.000700	0.001100
B6	1000.0031	0.003300	0.0859	0.000342	0.000408	0.008000	0.000739	0.001008
B7	1000.0072	0.009300	0.2863	0.001670	0.004080	0.000800	0.000740	0.001110

TABEL 2
DATA INTERKOMPARASI ARTEFAK 200 G

Kode	Mc 200g	Uc 200g	En	Ur	Ud	Ustd	Udrif	Ub
A1 (acuan)	200.00062	0.00057		0.0000415	0.0000041	0.00025	0.0001386	0.0000222
A2	200.0000	0.0082	0.0308	0.000000	0.004083	0.000250	0.000139	0.000222
A3	200.0030	0.010	0.3094	0.000408	0.000500	0.002475	0.000000	0.000222
A4	199.9999	0.0032	0.0455	0.000040	0.000040	0.001500	0.000460	0.000220
A5	200.00015	0.00067	0.3387	0.000000	0.000000	0.000300	0.000000	0.000200
A6	200.001	0.0010	0.8089	0.000000	0.000400	0.000200	0.000100	0.000200
A7	199.9993	0.0028	0.1276	0.000047	0.000041	0.000255	0.001390	0.000069
B1 (acuan)	200.00101	0.00057		0.000012	0.000004	0.000250	0.000139	0.000022
B2	200.00005	0.0083	0.0932	0.000000	0.004100	0.000300	0.000500	0.000200
B3	200.0015	0.0011	0.3743	0.000300	0.000400	0.000300	0.000000	0.069300
B4	200.0024	0.009200	0.1561	0.002134	0.004082	0.000050	0.000139	0.000222

Kode	Mc 200g	Uc 200g	En	Ur	Ud	Ustd	Udrif	Ub
B5	200.003	0.010	0.2010	0.002500	0.004100	0.000200	0.000100	0.000200
B6	200.00105	0.000740	0.1336	0.000052	0.000041	0.000250	0.000139	0.000222
B7	200.00114	0.000720	0.2054	0.000030	0.000040	0.000250	0.000140	0.000220

Artefak yang digunakan dalam kegiatan interkomparasi berupa artefak anak timbangan kelas M_1 dengan massa nominal 1 kg sebanyak dua unit yang dibedakan identitasnya dalam artefak 1 kg A dan artefak 1 kg B, serta artefak anak timbangan kelas M_2 dengan massa nominal 200 g sebanyak dua unit yang dibedakan identitasnya dalam artefak 200 g A dan artefak 200 g B.

Data interkomparasi pada tabel 1 dan tabel 2 terdiri dari:

A dan B yang merupakan kode peserta interkomparasi berdasarkan artefak yang digunakan. Tabel 1 dan tabel 2 menggunakan artefak A dan artefak B dengan data sebagai berikut:

Artefak A terdiri dari 2 anak timbangan:

1. Merek Ankatama, massa Nominal 1 kg, nomor Seri ATM.M-1. S.022, kelas M_1 , bahan stainless Steel dan massa jenis 7.950 kg/m³.
2. Tanpa merek, massa Nominal 200 g, tanpa no seri, kelas M_2 , bahan stainless Steel dan massa jenis 7.950 kg/m³

Artefak B terdiri dari 2 anak timbangan:

1. Merek Ankatama, massa Nominal 1 kg, nomor Seri ATM.M-1. S.021, kelas M_1 , bahan stainless Steel dan massa jenis 7.950 kg/m³.
2. Tanpa merek, massa Nominal 200 g, tanpa no seri, kelas M_2 , bahan stainless Steel dan massa jenis 7.950 kg/m³

M_c yang merupakan massa konvensional atau massa hasil penimbangan di udara pada kondisi konvensional yaitu suhu 20°C, massa jenis udara 1,2 kg/m³, dan massa jenis benda yang ditimbang sebesar 8000 kg/m³[16].

U_c merupakan ketidakpastian pengukuran yang menunjukkan nilai sebaran hasil pengukuran yang dikaitkan dengan besaran yang diukur [18].

En merupakan nilai rasio yang dihitung dari persamaan 5.

U_r merupakan nilai ketidakpastian yang bersumber dari *repeatability* timbangan elektronik yang digunakan.

U_d merupakan nilai ketidakpastian akibat dari daya baca / resolusi timbangan.

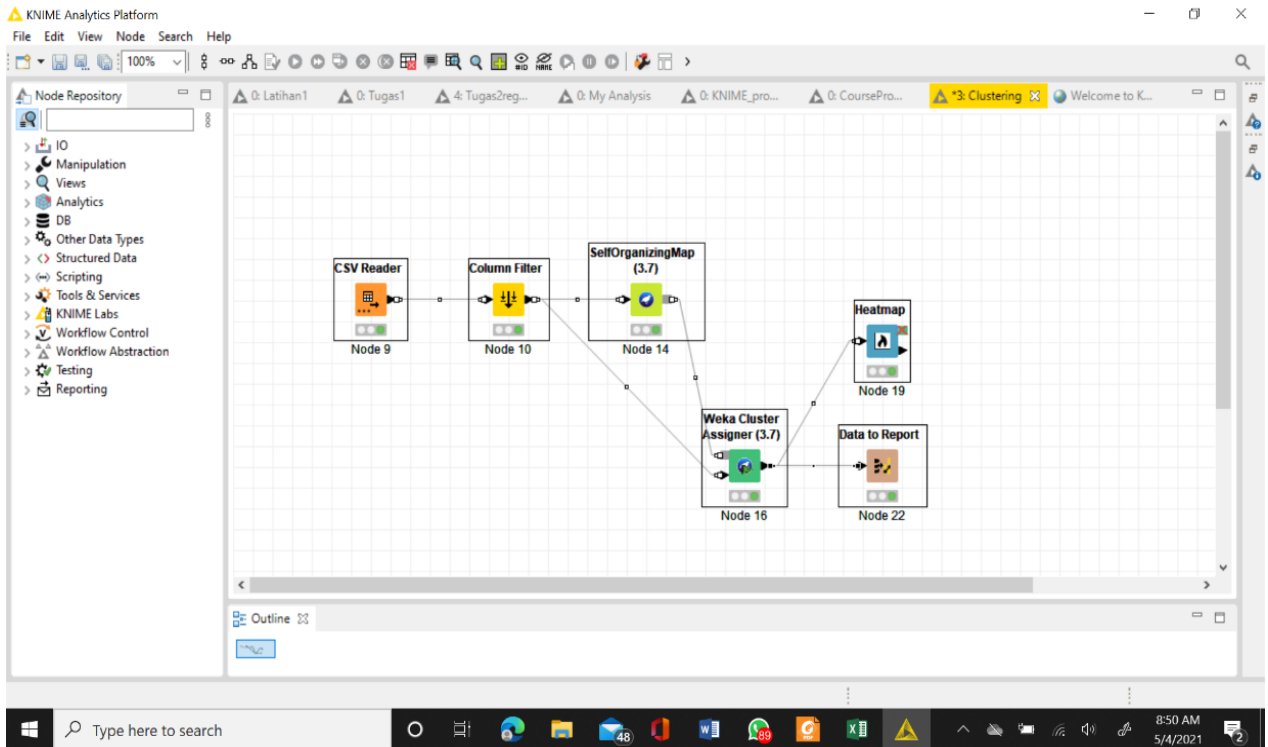
U_{std} merupakan nilai ketidakpastian yang bersumber dari penggunaan anak timbangan standar.

U_{drift} merupakan nilai ketidakpastian yang bersumber dari pergeseran nilai anak timbangan standar minimum dari 3 sertifikat terakhir

U_b merupakan nilai ketidakpastian yang bersumber dari koreksi gaya apung udara.

III. HASIL DAN PEMBAHASAN

Data interkomparasi dikelompokkan berdasar nilai En dan sumber ketidapastiannya dengan menggunakan platform "KNIME". KNIME atau *Konstanz Information Miner*, adalah platform analisis, pelaporan, dan integrasi data yang bersifat *open source* [14]. KNIME menyatukan berbagai komponen dalam bidang *machine learning* dan data *mining* dengan konsep *open data* yang modular. Dari gambar 3 proses clustering diawali dengan membaca data dengan format .csv pada *csv reader*, data tersebut kemudian dilakukan *filtering* pada *column filter* untuk menentukan atribut/variabel yang menjadi masukan pada proses *clustering*. Selanjutnya pada *Self Organizing Maps* dilakukan pengaturan nilai parameter dari *learning rate*, tinggi dan lebar *lattice*, jumlah *epoch* untuk fase *ordering* dan fase *convergence* dan status normalisasi data. Terdapat dua status normalisasi, *true* jika data dilakukan normalisasi dan *false* jika data tidak dilakukan normalisasi.



Gambar 3 Clustering dengan algoritma SOM menggunakan platform KNIME

Adapun pengaturan nilai *learning rate*, tinggi dan lebar *lattice* serta jumlah *epoch* untuk fase *ordering* dan fase *convergence* ditampilkan pada tabel 3 berikut:

TABEL 3
NILAI PARAMETER PERCOBAAN *CLUSTERING* DENGAN ALGORITMA SOM

Percobaan ke	learning rate	tinggi lattice	lebar lattice	Epoch _o	Epoch _c	Normalisasi
1	0,8	2	3	6000	2000	True
2	0,8	1	3	6000	2000	True
3	0,8	1	2	6000	2000	True
4	0,8	2	2	6000	2000	True
5	0,8	2	3	6000	2000	False
6	0,8	1	3	6000	2000	False
7	0,8	1	2	6000	2000	False
8	0,8	2	2	6000	2000	False

Data – data pada tabel 1 dan tabel 2 kemudian dikelompokkan dengan menggunakan algoritma *self organizing maps* sesuai dengan nilai parameter pada tabel 3. Nilai parameter pada percobaan ke-1 dan percobaan ke-5 sama hanya berbeda di status normalisasi datanya. Data yang dikelompokkan berupa nilai *En* dan *budget uncertainty* yang terdiri dari:

1. Ketidakpastian repeatability timbangan (U_r);
2. Ketidakpastian akibat daya baca / resolusi timbangan (U_d);
3. Ketidakpastian anak timbangan standar yang digunakan (U_{std});
4. Ketidakpastian *drift* atau pergeseran nilai dari anak timbangan standar (U_{drift});
5. Ketidakpastian akibat *bouyancy* udara (U_b).

Berikut hasil *Clustering* data interkomparasi artefak 1 kg menggunakan algoritma *self organizing maps* :

Row ID	S Kode	D En Ik	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	D Cluster3	D Cluster4	D Cluster5	I Winner ...
Kab Kediri	A2	1.297	0	0.004	0.001	0.001	0.001							0
Kab Demak	A3	0.775	0.004	0.002	0.011	0	0.001							5
Kab Semarang	A4	0.195	0.003	0.004	0.002	0.001	0.001							3
Kota Magelang	A5	0.116	0.001	0.004	0.001	0	0.001							3
Kota Cimahi	A6	5.59	0	0	0.001	0.001	0.001							2
Kota Semarang	A7	0.31	0	0.004	0.001	0.002	0							0
Kab Temangg...	B2	0.258	0	0.004	0.001	0.001	0.001							0
Kab Banyumas	B3	0.004	0.002	0.004	0.001	0	0.069							3
Kab Kuningan	B4	0.259	0.003	0.004	0	0.001	0.001							3
Kab Pekalongan	B5	0.584	0.003	0.004	0.001	0.001	0.001							3
Kab Bojonegoro	B6	0.086	0	0	0.008	0.001	0.001							2
Kab Kudus	B7	0.286	0.002	0.004	0.001	0.001	0.001							3

Gambar 4. Hasil *Clustering* SOM artefak 1 kg pada Percobaan ke-1 data dinormalisasi

Pada gambar 4 terdapat 6 *cluster* terbentuk dimana 2 *cluster* diantaranya tidak mempunyai anggota. *Cluster* yang tidak mempunyai anggota disebabkan oleh 2 hal, ukuran dimensi layer dan perlakuan datanya (dinormalisasi atau tidak). Tujuan data dinormalisasi adalah untuk menskalakan nilai atribut dari data sehingga bisa terbentuk data pada range tertentu di setiap atribut. Normalisasi data dilakukan dengan metode min-max dimana data pada atribut yang sama dibagi dengan nilai maksimum pada atribut tersebut. Data interkomparasi mayoritas memiliki rentang nilai yang relatif sama, pada gambar 4 terdapat 2 data yang outlier dari perhitungan En – nya yaitu A2 dan A6, sehingga data asli tidak perlu dilakukan normalisasi.

Row ID	S Kode	D En Ik	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	D Cluster3	D Cluster4	D Cluster5	I Winner ...
Kab Kediri	A2	1.297	0	0.004	0.001	0.001	0.001							2
Kab Demak	A3	0.775	0.004	0.002	0.011	0	0.001							4
Kab Semarang	A4	0.195	0.003	0.004	0.002	0.001	0.001							0
Kota Magelang	A5	0.116	0.001	0.004	0.001	0	0.001							1
Kota Cimahi	A6	5.59	0	0	0.001	0.001	0.001							5
Kota Semarang	A7	0.31	0	0.004	0.001	0.002	0							4
Kab Temangg...	B2	0.258	0	0.004	0.001	0.001	0.001							1
Kab Banyumas	B3	0.004	0.002	0.004	0.001	0	0.069							3
Kab Kuningan	B4	0.259	0.003	0.004	0	0.001	0.001							3
Kab Pekalongan	B5	0.584	0.003	0.004	0.001	0.001	0.001							4
Kab Bojonegoro	B6	0.086	0	0	0.008	0.001	0.001							1
Kab Kudus	B7	0.286	0.002	0.004	0.001	0.001	0.001							3

Gambar 5. Hasil *Clustering* SOM artefak 1 kg pada Percobaan ke-5 data tidak dinormalisasi

Hasil *clustering* dimana data interkomparasi tidak dinormalisasi ditunjukkan pada gambar 5, dimana terdapat 6 *cluster* terbentuk dan masing-masing *cluster* memiliki anggota. Anggota cluster seperti ditampilkan pada tabel 4, dimana masing-masing laboratorium dikelompokkan berdasar kedekatan nilai atribut yang dimiliki. Anggota *cluster* untuk percobaan ke-1 seperti ditampilkan pada tabel 5. Seperti pada cluster ke-2, ke-3 dan ke-4 yang masing-masing terdiri dari 3 anggota, masing-masing anggota memiliki kesamaan / kedekatan data. Demikian juga hasil percobaan yang ditunjukkan pada gambar 6 dan gambar 7.

TABEL 4
ANGGOTA CLUSTER PERCOBAAN KE-5

Nama Cluster	Anggota	En	Centroid	SSW	SSB
Cluster ke-1 (cluster0)	Kab. Semarang	0.195	0.195	0	0.095
Cluster ke-2 (cluster1)	Kota Magelang	0.116	0.153	0.017	0.093
	Kab. Temanggung	0.258			
	Kab. Bojonegoro	0.086			
Cluster ke-3 (cluster2)	Kab. Kediri	1.297	1.297	0	1.436
Cluster ke-4 (cluster3)	Kab. Banyumas	0.004	0.183	0.052	0.089
	Kab. Kuningan	0.259			
	Kab. Kudus	0.286			
Cluster ke-5 (cluster4)	Kab. Demak	0.775	0.5563	0.109	0.267
	Kota Semarang	0.31			
	Kab. Pekalongan	0.584			
Cluster ke-6 (cluster5)	Kota Cimahi	5.59	5.59	0	29.821

TABEL 5
ANGGOTA CLUSTER PERCOBAAN KE-1

Nama Cluster	Anggota	En	Centroid	SSW	SSB
Cluster ke-1 (cluster0)	Kab. Kediri	1.297	0.6217	0.47	0.162
	Kota Semarang	0.31			
	Kab. Temanggung	0.258			
Cluster ke-2 (cluster2)	Kota Cimahi	5.59	2.838	0.561	0.204
	Kab. Bojonegoro	0.086			
Cluster ke-3 (cluster3)	Kab. Semarang	0.195	0.241	1.712	0.151
	Kota Magelang	0.116			
	Kab. Banyumas	0.004			
	Kab. Kuningan	0.259			
	Kab. Pekalongan	0.584			
Cluster ke-4 (cluster5)	Kab. Kudus	0.286	0.775	0	1.584
	Kab. Demak	0.775			

Pada tabel 4 dan tabel 5 dari proses clustering data input yang terdiri dari 6 variabel antara lain En , U_c , U_r , U_d , U_{std} , U_{drift} , dan U_b masing-masing menyesuaikan diri di lingkungan grid / topologi yang sudah ditentukan nilai parameternya (ukuran layer 2x3). Terlihat pada tabel 4, klaster ke-1 diisi oleh kabupaten Semarang, sementara klaster ke-2 diisi oleh kota Magelang, kabupaten Temanggung dan kabupaten Bojonegoro, klaster ke-3 diisi oleh kabupaten Kediri, klaster ke-4 diisi oleh kabupaten Banyumas, kabupaten Kuningan dan kabupaten Kudus, klaster ke-5 diisi oleh kabupaten Demak, kota Semarang dan kabupaten Pekalongan, dan klaster ke-6 diisi oleh kota Cimahi. Masing-masing klaster memiliki nilai SSW (*sum square of within*) yang menggambarkan kedekatan antar anggota cluster dan nilai SSB (*sum square of between*) yang menggambarkan jarak antar cluster. Untuk data yang tidak dinormalisasi, kota Cimahi akan selalu berada di cluster terpisah, karena nilai En nya outlier dan jarak antar datanya cukup jauh. Berbeda pada data yang dinormalisasi dengan metode min – max, kota Cimahi bergabung dengan kabupaten Bojonegoro ke cluster 2 seperti tampak pada tabel 5.

Row ID	S Kode	D En 1k	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	I Winner ...
Kab Kediri	A2	1.297	0	0.004	0.001	0.001	0.001	█			0
Kab Demak	A3	0.775	0.004	0.002	0.011	0	0.001			█	2
Kab Semarang	A4	0.195	0.003	0.004	0.002	0.001	0.001	█			0
Kota Magelang	A5	0.116	0.001	0.004	0.001	0	0.001	█			0
Kota Cimahi	A6	5.59	0	0	0.001	0.001	0.001			█	2
Kota Semarang	A7	0.31	0	0.004	0.001	0.002	0	█			0
Kab Temangg...	B2	0.258	0	0.004	0.001	0.001	0.001	█			0
Kab Banyumas	B3	0.004	0.002	0.004	0.001	0	0.069	█			0
Kab Kuningan	B4	0.259	0.003	0.004	0	0.001	0.001	█			0
Kab Pekalongan	B5	0.584	0.003	0.004	0.001	0.001	0.001	█			0
Kab Bojonegoro	B6	0.086	0	0	0.008	0.001	0.001			█	2
Kab Kudus	B7	0.286	0.002	0.004	0.001	0.001	0.001	█			0

Gambar 6. Hasil Clustering SOM artefak 1 kg pada Percobaan ke-2 data dinormalisasi

Perubahan ukuran dimensi layer yang lebih kecil menghasilkan jumlah cluster yang lebih sedikit dibandingkan dengan ukuran dimensi layer yang lebih besar. Seperti tampak pada gambar 6 dan gambar 7 ukuran layernya diatur pada ketinggian 1 dan lebar 3

Row ID	S Kode	D En 1k	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	I Winner ...
Kab Kediri	A2	1.297	0	0.004	0.001	0.001	0.001	█			1
Kab Demak	A3	0.775	0.004	0.002	0.011	0	0.001			█	1
Kab Semarang	A4	0.195	0.003	0.004	0.002	0.001	0.001	█			0
Kota Magelang	A5	0.116	0.001	0.004	0.001	0	0.001	█			1
Kota Cimahi	A6	5.59	0	0	0.001	0.001	0.001			█	2
Kota Semarang	A7	0.31	0	0.004	0.001	0.002	0	█			0
Kab Temangg...	B2	0.258	0	0.004	0.001	0.001	0.001	█			0
Kab Banyumas	B3	0.004	0.002	0.004	0.001	0	0.069	█			0
Kab Kuningan	B4	0.259	0.003	0.004	0	0.001	0.001	█			0
Kab Pekalongan	B5	0.584	0.003	0.004	0.001	0.001	0.001	█			1
Kab Bojonegoro	B6	0.086	0	0	0.008	0.001	0.001			█	0
Kab Kudus	B7	0.286	0.002	0.004	0.001	0.001	0.001	█			0

Gambar 7. Hasil Clustering SOM artefak 1 kg pada percobaan ke 6 data tidak dinormalisasi

Dari percobaan yang sudah dilakukan kemudian dihitung nilai *pseudo F statistic* dengan menggunakan rumus (6) untuk mendapatkan jumlah *cluster* optimum dan nilai *icdrate* dengan menggunakan rumus (10) dan diperoleh hasil seperti pada tabel 6 berikut:

TABEL 6
PERHITUNGAN PSEUDO F STATISTIC ARTEFAK 1 KG

Percobaan ke	Normalisasi	Jumlah Cluster	Nilai Pseudo F	Nilai icdrate
1	True	6	0.420195284	0.745696732
2	True	3	1.083382204	0.874853676
3	True	2	2.407516008	0.874853676
4	True	4	0.681129734	0.863299004
5	False	6	114.6492548	0.07132414
6	False	3	93.16202824	0.114728143
7	False	2	113.3724626	0.124806003
8	False	4	167.531742	0.096350675

Dari perhitungan *pseudo F statistic* untuk artefak 1 kg diperoleh *cluster* optimum pada percobaan ke-8 dengan nilai *pseudo F* terbesar 167.53. Setting parameter pada percobaan ke-8 ini antara lain: nilai *learning rate* 0.8, tinggi *lattice* 2, lebar *lattice* 2, jumlah *epoch* fase order 6000, jumlah *epoch* fase *convergen* 2000, setting normalisasi *false* dan jumlah *cluster* optimum yang terbentuk sebanyak 4 *cluster* seperti tampak pada gambar 8 berikut ini.

Row ID	S Kode	D En Ik	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	D Cluster3	I Winner ...
Kab Kediri	A2	1.297	0	0.004	0.001	0.001	0.001					1
Kab Demak	A3	0.775	0.004	0.002	0.011	0	0.001					1
Kab Semarang	A4	0.195	0.003	0.004	0.002	0.001	0.001					3
Kota Magelang	A5	0.116	0.001	0.004	0.001	0	0.001					3
Kota Cimahi	A6	5.59	0	0	0.001	0.001	0.001					0
Kota Semarang	A7	0.31	0	0.004	0.001	0.002	0					3
Kab Temangg...	B2	0.258	0	0.004	0.001	0.001	0.001					2
Kab Banyumas	B3	0.004	0.002	0.004	0.001	0	0.069					3
Kab Kuningan	B4	0.259	0.003	0.004	0	0.001	0.001					2
Kab Pekalongan	B5	0.584	0.003	0.004	0.001	0.001	0.001					2
Kab Bojonegoro	B6	0.086	0	0	0.008	0.001	0.001					3
Kab Kudus	B7	0.286	0.002	0.004	0.001	0.001	0.001					2

Gambar 8. Hasil Clustering SOM artefak 1 kg pada percobaan ke 8 data tidak dinormalisasi

Sedangkan untuk nilai *icdrate* terkecil diperoleh pada percobaan ke-5 dengan nilai *icdrate* sebesar 0.0713. Setting parameter pada percobaan ke-5 ini antara lain: nilai *learning rate* 0.8, tinggi *lattice* 2, lebar *lattice* 3, jumlah *epoch* fase order 6000, jumlah *epoch* fase *convergen* 2000, setting normalisasi *false* dan jumlah *cluster* terbaik berjumlah 6 *cluster* seperti tampak pada gambar 9 berikut ini.

Row ID	S Kode	D En Ik	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	D Cluster3	D Cluster4	D Cluster5	I Winner ...
Kab Kediri	A2	1.297	0	0.004	0.001	0.001	0.001							2
Kab Demak	A3	0.775	0.004	0.002	0.011	0	0.001							4
Kab Semarang	A4	0.195	0.003	0.004	0.002	0.001	0.001							0
Kota Magelang	A5	0.116	0.001	0.004	0.001	0	0.001							1
Kota Cimahi	A6	5.59	0	0	0.001	0.001	0.001							5
Kota Semarang	A7	0.31	0	0.004	0.001	0.002	0							4
Kab Temangg...	B2	0.258	0	0.004	0.001	0.001	0.001							1
Kab Banyumas	B3	0.004	0.002	0.004	0.001	0	0.069							3
Kab Kuningan	B4	0.259	0.003	0.004	0	0.001	0.001							3
Kab Pekalongan	B5	0.584	0.003	0.004	0.001	0.001	0.001							4
Kab Bojonegoro	B6	0.086	0	0	0.008	0.001	0.001							1
Kab Kudus	B7	0.286	0.002	0.004	0.001	0.001	0.001							3

Gambar 9. Hasil Clustering SOM artefak 1 kg pada percobaan ke 5 data tidak dinormalisasi

Untuk artefak 200 g, nilai *pseudo F statistic* dan nilai *icdrate* seperti pada tabel 7 berikut ini:

TABEL 7
PERHITUNGAN PSEUDO F STATISTIC ARTEFAK 200 G

Percobaan ke	Normalisasi	Jumlah Cluster	Nilai Pseudo F	Nilai icdrate
1	True	6	2.531334764	0.706020386
2	True	3	2.43219426	0.916757618
3	True	2	5.404876134	0.916757618
4	True	4	1.689269142	0.995156222

5	False	6	76.77694972	0.288953334
6	False	3	28.38677472	0.353372631
7	False	2	18.41642895	0.808731814
8	False	4	104.8638168	0.993541629

Dari perhitungan *pseudo F statistic* untuk artefak 200 g diperoleh *cluster* optimum pada percobaan ke-8 dengan nilai *pseudo F* terbesar 104.86. Setting parameter pada percobaan ke-8 ini antara lain: nilai *learning rate* 0.8, tinggi *lattice* 2, lebar *lattice* 2, jumlah *epoch* fase order 6000, jumlah *epoch* fase *convergen* 2000, setting normalisasi *false* dan jumlah *cluster* optimum yang terbentuk sebanyak 4 *cluster* seperti tampak pada gambar 10 berikut ini.

Row ID	S Kode	D En 200	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	D Cluster3	I Winner ...
Kab Kediri	A2	0.031	0	0.004	0	0	0					1
Kab Demak	A3	0.309	0	0.001	0.002	0	0					3
Kab Semarang	A4	0.045	0	0	0.002	0	0					1
Kota Magelang	A5	0.339	0	0	0	0	0					3
Kota Cimahi	A6	0.809	0	0	0	0	0					2
Kota Semarang	A7	0.128	0	0	0	0.001	0					1
Kab Temang...	B2	0.093	0	0.004	0	0.001	0					1
Kab Banyumas	B3	0.374	0	0	0	0	0.069					3
Kab Kuningan	B4	0.156	0.002	0.004	0	0	0					0
Kab Pekalongan	B5	0.201	0.003	0.004	0	0	0					0
Kab Bojonegoro	B6	0.134	0	0	0	0	0					1
Kab Kudus	B7	0.205	0	0	0	0	0					0

Gambar 10. Hasil *Clustering* SOM artefak 200 g pada percobaan ke 8 data tidak dinormalisasi

Sedangkan untuk nilai *icdrate* terkecil diperoleh pada percobaan ke-5 dengan nilai *icdrate* sebesar 0.2889. Setting parameter pada percobaan ke-5 ini antara lain: nilai *learning rate* 0.8, tinggi *lattice* 2, lebar *lattice* 3, jumlah *epoch* fase order 6000, jumlah *epoch* fase *convergen* 2000, setting normalisasi *false* dan jumlah *cluster* terbaik berjumlah 5 *cluster* seperti tampak pada gambar 11 berikut ini.

Row ID	S Kode	D En 200	D Ur	D Ud	D Ustd	D Udrif	D Ub	D Cluster0	D Cluster1	D Cluster2	D Cluster3	D Cluster4	D Cluster5	I Winner ...
Kab Kediri	A2	0.031	0	0.004	0	0	0							2
Kab Demak	A3	0.309	0	0.001	0.002	0	0							5
Kab Semarang	A4	0.045	0	0	0.002	0	0							2
Kota Magelang	A5	0.339	0	0	0	0	0							5
Kota Cimahi	A6	0.809	0	0	0	0	0							0
Kota Semarang	A7	0.128	0	0	0	0.001	0							3
Kab Temang...	B2	0.093	0	0.004	0	0.001	0							2
Kab Banyumas	B3	0.374	0	0	0	0	0.069							1
Kab Kuningan	B4	0.156	0.002	0.004	0	0	0							3
Kab Pekalongan	B5	0.201	0.003	0.004	0	0	0							3
Kab Bojonegoro	B6	0.134	0	0	0	0	0							3
Kab Kudus	B7	0.205	0	0	0	0	0							3

Gambar 11. Hasil *Clustering* SOM artefak 200 g pada percobaan ke 5 data tidak dinormalisasi

IV. SIMPULAN

Setelah dilakukan analisis terhadap hasil *Clustering* data interkomparasi dengan algoritma *self organizing maps*, dapat disimpulkan bahwa Jumlah *cluster* optimum terjadi pada percobaan ke-8 dengan nilai *Pseudo F Statistic* 167.53 untuk artefak 1 kg dan nilai *Pseudo F Statistic* 104.86 untuk artefak 200 g dengan jumlah optimum sebanyak 4 *cluster*. *Cluster* terbaik terjadi pada percobaan ke-5 dengan nilai *icdrate* 0.0713 untuk artefak 1 kg dengan jumlah *cluster* terbaik sebanyak 6 *cluster* dan nilai *icdrate* 0.2889 untuk artefak 200 g dengan jumlah *cluster* terbaik sebanyak 5 *cluster*. Hasil *Clustering* algoritma *self organizing maps* pada ukuran layer 2 x 3 pada percobaan ke-5 memiliki nilai homogenitas dan jarak antar *cluster* terbaik, sedangkan *Clustering* dengan ukuran layer 2 x 2 pada percobaan ke-8 memiliki jumlah *cluster* optimum. Dari hasil penelitian ini identifikasi permasalahan laboratorium peserta interkomparasi dapat dianalisis secara tepat berdasarkan hasil *clustering* dengan menggunakan algoritma SOM dimana parameter nilai *learning rate* 0.8, ukuran *layer* 2x3 dan data tidak perlu dilakukan normalisasi.

DAFTAR PUSTAKA

- [1] *Conformity assessment - General requirements for proficiency testing*, ISO/IEC 17043:2010, 2010.
- [2] Edy Irwansyah, M Faisal, *Advanced Clustering Teori dan Aplikasi*, edisi 1, Deepublish, Yogyakarta, 2015
- [3] Hendayanti Ni Putu Nanik, Putri Gusti Ayu Made Arna, Nurhida Maulida, "Ketepatan Klasifikasi Penerima Beasiswa Stimik Stikom Bali Dengan Hybrid *Self organizing maps* Dan Algoritma K-Means", *Jurnal Varian* Vol.2 No.1, Oktober. 2018.
- [4] Purbasari I Y, Puspaningrum E Y, Putra A B S, "Using Self-Organizing Map (SOM) for *Clustering* and Visualization of New Students based on Grades", *Journal of Physics: Conference Series* 1569 02203 IOP Publishing, 2020 Tersedia: <https://doi:10.1088/1742-6596/1569/2/022037>

- [5] Kania Siti, Rachmatin Dewi, Dahlan Jarnawi Afgani, "Program Aplikasi Pengelompokan Objek Dengan Metode Self Organizing Map Menggunakan Bahasa R", Jurnal EurekaMatika, Vol.7 No.2, hlm. 17-29. 2019
- [6] Firmansyah Dese Narfa; Adinugroho Sigit; Rahayudi Bayu, "Pengelompokan Wilayah Berdasarkan Kesejahteraan Sosial Menggunakan Algoritme Self-Organizing Maps Dengan Perbaikan Missing Value K-Nearest Neighbors", Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer Universitas Brawijaya Vol. 3, No. 7, hlm. 7205-7213, Juli 2019
- [7] Lakshminarayanan Sivakkumaran, "Application of Self-Organizing Maps on Time Series Data for identifying interpretable Driving Manoeuvres", SpringerOpen Jurnal, 2020. Tersedia: <https://doi.org/s12544-020-00421-x>
- [8] TIAN Jin, GU Mengyi, "Subspace Clustering Based on Self-Organizing Map", ResearchGate, 2019. Tersedia: [https:// DOI: 10.1007/978-981-13-3402-3_17](https://doi.org/10.1007/978-981-13-3402-3_17).
- [9] Negnevitsky Michael, "Artificial Intelligence: A Guide to Intelligent System", Third edition, Pearson Education Limited, England, 2011
- [10] Putri Marina M, "Pengelompokan Kabupaten/Kota Di Provinsi Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat Menggunakan Metode Kohonen Som Dan K-Means", Jurnal Sains dan Seni ITS Vol 4 No.1, 2015 Tersedia: [https:// DOI: 10.12962/j23373520.v4i1.8815](https://doi.org/10.12962/j23373520.v4i1.8815)
- [11] Cottrel M et all, "Self-Organizing Maps, Theory and Applications", Revista Investigacion Operacional VOL. 39, NO. 1, pp 1-22, 2018
- [12] Yin Hujun, "The Self-Organizing Maps: Background, Theories, Extensions and Applications", <https://www.researchgate.net/publication/227313166> DOI: 10.1007/978-3-540-78293-3_17, May 2008
- [13] Voiculescu, R. M., Olteanu, M. C., & Nistor, V. M. "Design and operation of an interlaboratory comparison scheme". Proceedings of NUCLEAR, (p. 159), 2013
- [14] M. R. Berthold, N. Cebon, F. Dill, T. R. Gabriel, T. Kötter, T. Meinl, P. Ohl, K. Thiel, B. Wiswedel, "KNIME - the Konstanz information miner", ACM SIGKDD Explorations Newsletter, vol 11, no 1, pp 26–31, 16 November 2009. doi:10.1145/1656274.1656280.
- [15] Program Interkomparasi Standar Massa - Protokol Interkomparasi BSML Regional II, 2021
- [16] Syarat Teknis Standar Ukuran Besaran Massa, Keputusan Direktur Jenderal Perlindungan Konsumen dan Tertib Niaga No 123 tahun 2020, 2020