

Nama:Kevin Christy Parinussa

NRP:1672064

Soal A

Pada gambar dibawah ini, kita meng-assign variabel start,end, stratIC, dan endIC untuk menampung nilai centroid. Setelah itu kita membuat sebuah variabel Configuration yang merupakan tempat dimana parameter K-means berada. Terdapat parameter K-means yang akan diambil dari config.xml, yaitu:

- DATASET_SIZE adalah jumlah poin dalam sebuah dataset
- DISTANCE adalah jarak yang digunakan
- K adalah jumlah partisi yang sudah ditemukan
- THRESHOLD adalah ambang pemberhentian kriteria
- MAX_ITERATIONS adalah jumlah maksimum pengulangan

Parameter tersebut diambil dari config.xml dimana variable input dan output dijalankan melalui command line.

```

long start = 0;
long end = 0;
long startIC = 0;
long endIC = 0;

start = System.currentTimeMillis();
Configuration conf = new Configuration();
conf.addResource(new Path("config.xml")); //Configuration file for the parameters

String[] otherArgs = new GenericOptionsParser(conf, args).getRemainingArgs();

if (otherArgs.length != 2) {
    System.err.println("Usage: <input> <output>");
    System.exit(1);
}

//Parameters setting
final String INPUT = otherArgs[0];
final String OUTPUT = otherArgs[1] + "/temp";
final int DATASET_SIZE = conf.getInt("dataset", 10);
final int DISTANCE = conf.getInt("distance", 2);
final int K = conf.getInt("k", 3);
final float THRESHOLD = conf.getFloat("threshold", 0.0001f);
final int MAX_ITERATIONS = conf.getInt("max.iteration", 30);

```

Pada gambar dibawah ini, kita membuat 2 variabel yaitu oldCentroids dan newCentroids untuk menyimpan himpunan data poin. Kemudian program akan menginisiasi centroid melalui loop, centroid baru dimasukkan kedalam config.xml. Setelahnya terdapat loop while yang menjalankan fungsi Map Reduce.

```

Point[] oldCentroids = new Point[K];
Point[] newCentroids = new Point[K];

//Initial centroids
startIC = System.currentTimeMillis();
newCentroids = centroidsInit(conf, INPUT, K, DATASET_SIZE);
endIC = System.currentTimeMillis();
for(int i = 0; i < K; i++) {
    conf.set("centroid." + i, newCentroids[i].toString());
}

//MapReduce workflow
boolean stop = false;
boolean succeeded = true;
int i = 0;
while(!stop) {
    i++;

    //Job configuration
    Job iteration = Job.getInstance(conf, "iter_" + i);
    iteration.setJarByClass(KMeans.class);
    iteration.setMapperClass(KMeansMapper.class);
    iteration.setCombinerClass(KMeansCombiner.class);
    iteration.setReducerClass(KMeansReducer.class);
    iteration.setNumReduceTasks(K); //one task each centroid
    iteration.setOutputKeyClass(IntWritable.class);
    iteration.setOutputValueClass(Point.class);
    FileInputFormat.addInputPath(iteration, new Path(INPUT));
    FileOutputFormat.setOutputPath(iteration, new Path(OUTPUT));
    iteration.setInputFormatClass(TextInputFormat.class);
    iteration.setOutputFormatClass(TextOutputFormat.class);
}

```

Jika semua system sudah berhasil dijalankan, maka sistem akan melakukan print waktu eksekusi program dan centroid.

```

        succeeded = iteration.waitForCompletion(true);

        //If the job fails the application will be closed.
        if(!succeeded) {
            System.err.println("Iteration" + i + "failed.");
            System.exit(1);
        }

        //Save old centroids and read new centroids
        for(int id = 0; id < K; id++) {
            oldCentroids[id] = Point.copy(newCentroids[id]);
        }
        newCentroids = readCentroids(conf, K, OUTPUT);

        //Check if centroids are changed
        stop = stoppingCriterion(oldCentroids, newCentroids, DISTANCE, THRESHOLD);

        if(stop || i == (MAX_ITERATIONS - 1)) {
            finalize(conf, newCentroids, otherArgs[1]);
        } else {
            //Set the new centroids in the configuration
            for(int d = 0; d < K; d++) {
                conf.unset("centroid." + d);
                conf.set("centroid." + d, newCentroids[d].toString());
            }
        }
    }

    end = System.currentTimeMillis();

    end -= start;
    endIC -= startIC;

```

```

        System.out.println("execution time: " + end + " ms");
        System.out.println("init centroid execution: " + endIC + " ms");
        System.out.println("n_iter: " + i);

        System.exit(0);
    }

```

Soal B

1.

Kebutuhan Informasi	Kueri	Hasil Mesin	Hasil Relevansi
Mobil terbaik 2021 di indonesia	Mobil terbaik 2021 di Indonesia About 138,000,000 results (0.62 seconds)	<u>10 SUV Terbaik & Terfavorit di Indonesia Terbaru 2021 Carro ...</u> https://carro.id › blog › beli-mobil › 10-suv-terbaik-terfav...	T
		<u>Mobil Paling Hot 2021 10 Mobil Terhebat Di Indonesia Oto</u> https://www.oto.com › mobil-populer	T
		<u>10 MPV Terbaik & Terfavorit di Indonesia Terbaru 2021 ...</u> https://carro.id › blog › beli-mobil	T
		<u>15 Mobil Keluarga Terbaik 2021: MPV Murah hingga Mewah</u> https://www.tokopedia.com › blog	T
		<u>10 Mobil Terbaik di Indonesia 2021 Ulasmobil</u> https://ulasmobil.com › 6 Pilihan	T
		<u>Daftar Mobil Terbaik di Indonesia 2021 - Wefixcar - Wefixit</u> https://wefixit.id › car › daftar-mobil...	T

		<u>Ini Rekomendasi Mobil Keluarga Terbaik di Tahun 2021 ...</u> https://www.goala.app > ... > Otomotif	T
		<u>10 Mobil Terlaris di Indonesia Maret 2021, Avanza Kembali ...</u> https://otomotif.kompas.com > read	T
		<u>√ Mobil Terbaik Indonesia 2021 Dengan Harga Termurah ...</u> https://www.otomotif.com > mobil-t...	T
		<u>100 Mobil SUV Ternyaman dan Terbaik Harga Murah 2021 ...</u> https://www.otoflik.com > mobil-suv...	T

2.

Kebutuhan Informasi	Perubahan kata pencarian	Kueri	Hasil Mesin	Hasil Relevansi
Mobil terbaik 2021 di Indonesia	Menghilangkan kata "di"	Mobil terbaik 2021 Indonesia About 99,200,000 results (0.58 seconds)	<u>10 SUV Terbaik & Terfavorit di Indonesia Terbaru 2021 Carro ...</u> https://carro.id > blog > beli-mobil > 10-suv-terbaik-terfav...	T
			<u>Mobil Paling Hot 2021 10 Mobil Terhebat Di Indonesia Oto</u> https://www.oto.com > mobil-populer	T
			<u>10 MPV Terbaik & Terfavorit di Indonesia Terbaru 2021 ...</u>	T

			https://carro.id › blog › beli-mobil	
			<u>5 Pilihan Mobil Baru Terbaik di Indonesia Selama Paruh ...</u> https://www.autofun.co.id › berita	T
			<u>15 Mobil Keluarga Terbaik 2021: MPV Murah hingga Mewah</u> https://www.tokopedia.com › blog	T
			11 Mobil SUV Terbaik 2021: Kini Lebih Terjangkau - Tokopedia https://www.tokopedia.com › blog	T
			<u>10 Mobil Terlaris di Indonesia Maret 2021, Avanza Kembali ...</u> https://otomotif.kompas.com › read	T
			√ <u>Mobil Terbaik Indonesia 2021 Dengan Harga Termurah ...</u> https://www.otomotif.com › mobil-t...	T
			Ini Rekomendasi Mobil Keluarga Terbaik di Tahun 2021 ... https://www.qoala.app › ... › Otomotif	T
			<u>Mobil Baru 2021 - Daftar Harga, Spesifikasi, Review & Promo ...</u> https://www.carmudi.co.id › mobilbaru	F

3.

Kebutuhan Informasi	Operator yang ditambahkan	Kueri	Hasil Mesin	Hasil Relevansi
Mobil terbaik 2021 di Indonesia	Menambahkan underscore	Mobil_terbaik_2021_di_Indonesia 10 results (0.53 seconds)	Mobil Paling Hot 2021 10 Mobil Terhebat Di Indonesia Oto https://www.oto.com › mobil-populer	T
			5 Asuransi Kendaraan Mobil Terbaik 2021 di Indonesia ... https://www.moneyfazz.id › asuransi-...	F
			Simak 10 Asuransi mobil terbaik 2021 di Indonesia Wajib ... https://pakgalingging.blogspot.com › ...	F
			Cara Mudah Service Injeksi Motor atau mobil EFI - Pak ... https://pakgalingging.blogspot.com › ...	F
			10 SUV Terbaik & Terfavorit di Indonesia Terbaru 2021 Carro ... https://carro.id › blog › beli-mobil	T
			Simak 10 Asuransi Kendaraan Beroda Empat Terbaik 2021 Di ... https://kau-menemukanmmu.blogspot.com › ...	F

			<u>Redaksi - Forum UMKM IKM Cibungbulang Bogor</u> https://umkmikmcibungbulang.com > ...	F
			<u>Travelling Archives - Forum UMKM IKM Cibungbulang Bogor</u> https://umkmikmcibungbulang.com > ...	F
			<u>Insurance – My Blog - Modekipsi</u> https://modekipsi.com > category > i...	F
			<u>Asuransi Mobil 2021 Money Nx - Melex.ID</u> https://www.melex.id > asuransi-mob...	F

4.

- a. Hasil temu pada Google jauh lebih besar dibandingkan Bing. Tidak, karena Google lebih banyak dipakai oleh banyak pengguna.
- b. Tidak selalu, karena ada beberapa hasil temu yang sangat memerlukan keyword yang sangat spesifik.
- c. Ya, tetapi tidak direkomendasikan, karena hasil temu akan menghasilkan hasil yang berbeda.
- d. Ya, karena search engine akan menghitung berapa jumlah keyword yang kita masukkan dan akan menunjukkan hasil yang relevan jika kita memasukkan keyword sebelumnya.
- e. Memasukkan kata dalam kueri dengan kalimat yang lengkap dan jelas agar hasil temu yang didapat akan lebih relevan.
- f. Google. Karena lebih banyak fitur dan hasil yang diberikan jauh lebih spesifik dibandingkan yang lain.