PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PEMETAAN WILAYAH POTENSIAL BERMASALAH NASABAH SEPEDA MOTOR

**Adrian Ahmad, Haditsah Annur, M.Kom1, Serwin, M.Kom2**

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Ichsan Gorontalo

Gorontalo, Indonesia

[*adrianahmad078@gmail.com*](mailto:adrianahmad078@gmail.com)

***Abstrak*** - Penerapan Algoritma Linear Regresi Untuk Memprediksi Persedian Obat Cair. Studi Kasus: RSUD. dr. Hasri Ainun Habibie. Penelitian ini bertujuan untuk menguji coba algoritma linear regresi untuk memprediksi jumlah sediaan obat cair pada bulan berikutnya, memperoleh sistem prediksi jumlah sediaan obat cair menggunakan algoritma linear regresi. Penelitian ini dilaksanakan di RSUD dr. Hasri Ainun Habibie dengan objek penelitian persediaan obat cair. Metode pengumpulan data yang digunakan dalam penelitian adalah teori-teori tentang prediksi, metode linear regresi maupun tentang obat cair. pengembangan model menggunakan metode linear regresi kemudian di evaluasi dengan menggunakan mean absolute persentase error. Konstruksi sistem menggunakan alat bantu PHP (Personal Home Page) dengan sistem database. Tahap pengujian dalam penelitian yaitu white box testing dan black box testing, kemudian dilanjutkan dengan pengujian terakhir yaitu user acceptance testing. Hasil penelitian menunjukkan bahwa prediksi persediaan obat cair RSUD. dr. Hasri Ainun Habibie menggunakan algoritma linier regresi sederhana yang di uji kinerjanya dengan metode white box testing menghasilkan V(G)=CC = 12, sehingga dinyatakan bahwa sistem ini telah memenuhi syarat logika pemrograman dan tidak kompleks. Sedangkan pengujian black box testing menyatakan bahwa sistem ini telah bebas dari berbagai kesalahan komponen-komponennya.

Kata Kunci: Linear Regresi, Prediksi Persedian Obat Cair

**I. PENDAHULUAN**

Berdasarkan ruang lingkup usahanya, Adira Finance dapat diklasifikasikan sebagai perusahaan pembiayaan non-bank dalam bidang pembiayaan konsumen. Obyek pembiayaan Adira Finance yaitu kendaraan bermotor roda dua dan empat. Terdapat dua macam target pasar Adira Finance, yaitu nasabah individu dan perusahaan. Segmen target individu Adira Finance adalah masyarakat kelas menengah ke bawah yang membutuhkan pembiayaan kendaraan bermotor roda dua (sepeda motor), sedangkan untuk jasa pembiayaan mobil, Adira Finance menargetkan masyarakat kalangan menengah ke atas. Dalam menjalankan operasional bisnis, Adira Finance berkaitan dengan pihak lain yaitu Dealer/Showroom dan Konsumen/Nasabah [1]. Dalam dunia ilmu komputer, dikenal luas data mining sebagai teknik penggalian data untuk mencari sebuah pola tersembunyi demi menghasilkan sebuah pengetahuan baru di dalam sekumpulan data [3]. Secara khusus data mining memiliki metode tersendiri berdasarkan tujuan dari pemanfaatan himpunan data yaitu estimasi, prediksi, klasifikasi, klasterisasi, dan asosiasi.

Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster. Objek yang di dalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma clustering. Oleh karena itu, clustering sangat berguna dan bisa menemukan group atau kelompok yang tidak dikenal dalam data. Clustering banyak digunakan dalam berbagai aplikasi seperti misalnya pada business inteligence, pengenalan pola citra, web search, bidang ilmu biologi, dan untuk keamanan (security). Hasil clustering yang baik akan menghasilkan tingkat kesamaan yang tinggi dalam satu kelas dan tingkat kesamaan yang rendah antar kelas. Kesamaan yang dimaksud merupakan pengukuran secara numeric terhadap dua buah objek. Nilai kesamaan antar kedua objek akan semakin tinggi jika kedua objek yang dibandingkan memiliki kemiripan yang tinggi. Begitu juga dengan sebaliknya. Kualitas hasil clustering sangat bergantung pada metode yang dipakai [4].

Metode K-means merupakan metode clustering yang paling sederhana dan umum. Hal ini dikarenakan K-means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien. K-Means merupakan salah satu algoritma klastering dengan metode partisi (partitioning method) yang berbasis titik pusat (centroid) selain algoritma k-Medoids yang berbasis obyek. Algoritma ini pertama kali diusulkan oleh MacQueen (1967) dan dikembangkan oleh Hartigan dan Wong tahun 1975 dengan tujuan untuk dapat membagi M data point dalam N dimensi kedalam sejumlah k cluster dimana proses klastering dilakukan dengan meminimalkan jarak sum squares antara data dengan masing masing pusat cluster (centroid-based). Algoritma k-Means dalam penerapannya memerlukan tiga parameter yang seluruhnya ditentukan pengguna yaitu jumlah cluster k, inisialisasi klaster, dan jarak system, Biasanya, k-Means dijalankan secara independen dengan inisialisasi yang berbeda menghasilkan cluster akhir yang berbeda karena algoritma ini secara prinsip hanya mengelompokan data menuju local minimal. Salah satu cara untuk mengatasi local minima adalah dengan mengimplementasikan algoritma k-Means, untuk K yang diberikan, dengan beberapa nilai initial partisi yang berbeda dan selanjutnya dipilih partisi dengan kesalahan kuadrat terkecil [4].

Berdasarkan pemaparan diatas, maka dapat dirumuskan masalah penelitian sebagai berikut: (1) “Bagaimana klasterisasi menggunakan metode K-Means untuk memetakan wilayah berpotensi bermasalah nasabah sepeda motor PT. Adira Finance Cabang Kota Gorontalo?”; (2) “Bagaimana tingkat akurasi terhadap hasil penerapan klasterisasi menggunakan metode K-Means untuk memetakan wilayah berpotensi bermasalah nasabah sepeda motor PT. Adira Finance Cabang Kota Gorontalo?”. Penelitian ini bertujuan untuk : (1) “Untuk mengklasterisasi menggunakan metode K-Means untuk memetakan wilayah berpotensi bermasalah nasabah sepeda motor PT. Adira Finance Cabang Kota Gorontalo; (2) “Untuk mengetahui tingkat akurasi hasil penerapan klasterisasi menggunakan metode K-Means untuk memetakan wilayah berpotensi bermasalah nasabah sepeda motor PT. Adira Finance Cabang Kota Gorontalo.”. Diharapkan apabila tujuan penelitian ini tercapai, maka Penelitian ini diharapkan dapat memberikan solusi bagi peneliti lain yang akan melakukan penelitian lanjutan untuk menerapkan dalam sistem yang lebih kompleks dan memberikan masukan bagi perkembangan ilmu pengetahuan dan teknologi, khususnya pada bidang ilmu komputer serta Memberikan sumbangan pemikiran, karya, bahan pertimbagan, atau solusi bagi software developer guna mendukung pengambilan keputusan dalam rangka menghasilkan software yang berkualitas sehingga berdampak pada peningkatan kualitas perusahaan.

Sistematika penyusunan artikel ilmiah ini pada bagian 2 menjelaskan tentang penelitian terkait, bagian 3 menjelaskan tentang tahapan penelitian, bagian 4 model penelitian, bagian 5 *sofware development*, bagian 6 hasil dan pembahasan, bagian 7 kesimpulan.

**II. PENELITIAN TERKAIT**

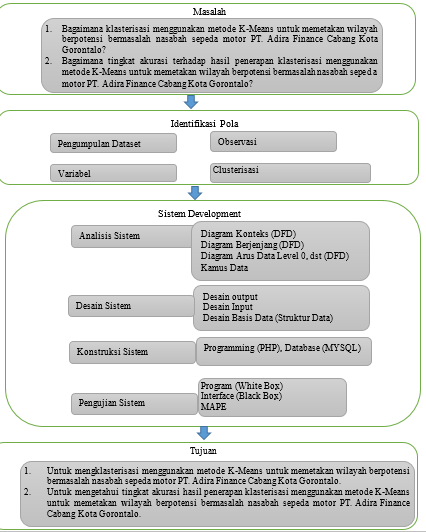
Penelitian terkait ditunjukan pada Tabel 1 berikut ini.

Tabel 1: Penelitian Terkait

|  |  |  |
| --- | --- | --- |
| NO | PENELITIAN TERKAIT | |
| 1 | PENELITI | Windha Mega Pradnya Dhuhita |
| JUDUL | Clustering menggunakan Metode K-Means untuk Menentukan Status Gizi Balita. |
| TAHUN | 2015 |
| METODE | K-Means |
| HASIL | 34% |
| 2 | PENELITI | Nurul Rohmawati W dkk |
| JUDUL | Implementasi Algoritma K-Means dalam Pengklasteran Mahasiswa Pelamar Beasiswa. |
| TAHUN | 2015 |
| METODE | K-Means |
| HASIL | 75% |
| 3 | PENELITI | Anindya Khrisna Wardhani |
| JUDUL | *K-Means Algorithm Implementation for Clustering of Patients Disease in* *Kajen Clinic of Pekalongan.* |
| TAHUN | 2016 |
| METODE | K-Means |
| HASIL | *This* *research used tool Rapid Miner 5.3.Based on the data from clinic centers Kajen Pekalongan, the result of clustering is 376 items of acute and 624* *unacute diseases from 1000 total of data*. |

**III. TAHAPAN PENELITIAN**

Tahapan penelitian ditunjukan pada Gambar 1 berikut ini.



Gambar 1: Tahapan Penelitian

**IV. PEMODELAN**

1. **Pengumpulan Data**

Untuk mengumpulkan data digunakan 2 (dua) jenis data yaitu data primer dan data sekunder. Data primer yaitu data yang berasal dari penelitian lapangan dan data sekunder berasal dari penelitian kepustakaan.

1. Penelitian Data Primer (Lapangan)Untuk memperoleh data primer yang merupakan data langsung dari objek penelitian yaitu bertempat di PT. Adira Finance Cabang Kota Gorontalo. Maka dilakukan dengan teknik:
2. Observasi, metode ini memungkinkan analis sistem mengamati atau meninjau langsung. Adapun pada penelitian ini dilakukan dengan cara menggumpulkan data wilayah berpotensi bermasalah nasabah sepeda motor di PT. Adira Finance Cabang Kota Gorontalo.
3. Wawancara metode ini digunakan dengan mangajukan beberapa pertanyaan kepada bagian yang terkait pada PT. Adira Finance Cabang Kota Gorontalo. Lingkup penelitian wilayah berpotensi bermasalah nasabah sepeda motor.
4. Penelitian Data Sekunder (Kepustakaan)

Metode kepustakaan diperlukan untuk mendapatkan data sekunder dengan tujuan melengkapi data primer. Data sekunder didapatkan dari pengkajian kepustakaan yang berisi dasar-dasar teori. Metode kepustakaan digunakan oleh analis sistem dengan cara mengambil contoh dokumen-dokumen yang berhubungan dengan materi penelitian. Selain itu, analis sistem mencari data mengenai hal-hal atau parameter yang berupa catatan, buku, majalah, dan sebagainya yang berhubungan dengan penelitian.

1. Pengembangan Model

Prosedur atau langkah-langkah pokok dalam klasterisasi menggunakan metode K-Means untuk memetakan wilayah berpotensi bermasalah nasabah sepeda dengan menggunakan alat bantu tools PHP, Database MySQL serta White Box Testing dan Black Box Testing untuk menguji kinerja sistemnya.

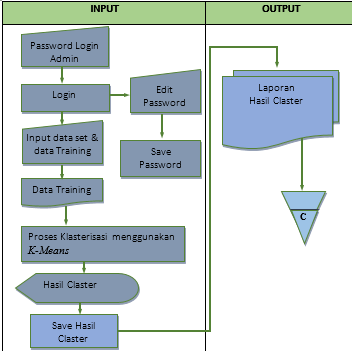
1. Evaluasi Model

Model yang telah dihasilkan kemudian dievaluasi dengan menggunakan MAPE untuk mengetahui Error.

1. **Pengembangan Sistem**

Sistem yang diusulkan dapat digambarkan menggunakan flowchart berikut ini

Gambar 2 : Pengembangan Sistem



1. Analisa Sistem

Analisis sistem menggunakan pendekatan berorientasi procedural/structural:

* Diagram Konteks, menggunanakan alat bantu DFD
* Diagram Berjenjang, menggunanakan alat bantu DFD
* Diagram Arus Data Level 0,1, dst. menggunakan alat bantu DFD
* Kamus Data menggunakan alat bantu Ms. Word.

1. Desain Sistem

Desain sistem menggunakan pendekatan berorientasi objek yang digambarkan dalam bentuk :

* Desain Output

Desain output dimaksudkan untuk mengetahui bagaimana dan seperti apa bentuk output-output dari sistem yang akan dibuat. Desain output terinci terbagi atas dua, yaitu desain output berbentuk laporan di media kertas dan desain output dalam bentuk dialog di layar terminal (monitor).

* Desain Input

Masukan merupakan awal dimulainya proses pengolahan informasi. Bahan mentah dari informasi adalah data yang terjadi dari transaksi-transaksi yang dilakukan oleh konsumen. Data hasil dari transaksi tidak lepas dari data yang dimasukkan. Desain input terinci dimulai dari desain dokumen dasar sebagai penangkap input yang pertamakali. Jika dokumen dasar tidak didesain dengan baik, kemungkinan input yang tercatat dapat salah bahkan kurang.

* Desain Database

Basis data (database) merupakan kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan di simpanan luar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. Database merupakan salah satu komponen yang penting di sistem informasi, karena berfungsi sebagai basis penyedia informasi bagi para pemakainya. Penerapan database dalam aplikasi disebut database system.

* Desain Teknologi

Pada tahap ini kita menentukan teknologi yang akan dipergunakan dalam menerima input, menjalankan model, menyimpan dan mengakses data, menghasilkan dan mengirimkan keluaran dan membantu pengendalian dari sistem secara keseluruhan.

* Desain Program

Pada tahap ini menggunakan alat bantu PHP dalam bentuk Pseudoce program pada untuk proses klasterisasi menggunakan metode K-Means untuk klasterisasi wilayah berpotensi bermasalah nasabah sepeda menggunakan algoritma K-Means.

1. Konstruki Sistem

Pada tahap ini dilakukan pembuatan sistem menggunakan tools PHP dan Database MySQL serta White Box Testing dan Black Box Testing untuk menguji kinerja sistem dan pengukuran akurasi menggunakan MAPE. Pada tahap ini kita melakukan tahap produksi sistem hasil analisa dan desain sistem sebelumnya. Termasuk didalamnya menginstal paket tambahan untuk menjalankan program, menulis listing program dan membangunnya dalam bentuk sebuah formulir, antar muka dan integrasi sistem-sistem program yang terdiri dari input, proses dan output yang tersusun dalam sebuah sistem menu sehingga dapat dijalankan oleh pengguna sistem.

1. Pengujian Sistem

Setelah dilakukan tahap analisa, desain dan produksi sistem, maka kita melakukan tahap pengujian, dimana seluruh perangkat lunak, program tambahan dan semua program yang terlibat dalam pembangunan sistem diuji untuk memastikan sistem dapat berjalan dengan semestinya. Testing difokuskan pada logika internal, fungsi eksternal dan mencari segala kemungkinan kesalahan dari sistem yang dibuat. Pada tahap ini dilakukan review dan evaluasi terhadap sistem yang dikembangkan, apakah sudah sesuai dengan rancangan atau belum. Jika terjadi hal-hal yang tidak sesuai dengan yang diharapkan, kemudian dilakukan revisi atau perbaikan supaya produk tersebut dapat dioperasikan dengan baik dan siap untuk diimplementasikan. Pengujian yang dilakukan dengan menggunakan teknik pengujian perangkat lunak yaitu:

* Pengujian White Box
* Pengujian Black Box

**V. HASIL PEMODELAN**

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Kota/Kecamatan** | **Waktu** | **Jumlah Nasabah** |
| 1 | Bongomeme | 22790 | 28 |
| 2 | Batudaa | 22699 | 20 |
| 3 | Atinggola | 18727 | 5 |
| 4 | Sumalata | 19649 | 13 |
| 5 | Telaga | 17714 | 7 |
| 6 | Limboto | 11627 | 12 |
| 7 | Isimu | 0 | 0 |
| 8 | Randangan | 90 | 1 |
| 9 | Marisa | 22885 | 18 |
| 10 | Paguat | 20114 | 5 |
| 11 | Paguyaman | 22692 | 27 |
| 12 | Kwandang | 19655 | 6 |
| 13 | Tapa | 20576 | 6 |
| 14 | Kabila | 93 | 1 |
| 15 | Suwawa | 22878 | 23 |
| 16 | Kota Tengah | 20570 | 21 |
| 17 | Kota Barat | 19469 | 6 |
| 18 | Kota Selatan | 19281 | 38 |
| 19 | Kota Utara | 19651 | 12 |

Tabel 2: Hasil Pengumpulan Data

1. Pseudocode Algoritma K-Means Untuk Cluster Pemetaan Wilayah

* Inisialisasi: memilih objek k secara acak yang akan berfungsi sebagai means.
* Mengasosiasikan setiap titik data dengan means yang paling serupa dengan mengguna kan ukuran jarak dan menghitung biaya.
* Secara acak memilih objek k baru yang akan berfungsi sebagai means dan menyimpan salinan dari set asli.
* Gunakan set means baru untuk menghitung ulang biaya.
* Jika biaya yang baru lebih besar dari pada biaya lama kemudian hentikan algoritma tersebut.
* Ulangi langkah kedua hingga kelima sampai tidak ada perubahan dalam means.

**VI. HASIL DAN PEMBAHASAN**

1. Pembahasan Model

Pembahasan dari cluster untuk pemetaan wilayah potensial bermasalah nasabah motor yang telah dihasilkan, dapat berupa :

Pada proses pemetaan produksi pada penelitian ini diawali dengan penginputan data yang didapatkan sebelumnya, dari data tersebut akan dibentuk centroid awal secara random dan dari centroid itu akan dibentuk cluster.

1. Pembahasan Sistem

Pembahasan sistem dapat berupa :

* Pembahasan terkait hasil pengujian sistemnya

Berdasarkan pengujian sistem pada data yang digunakan yaitu pada jumlah jumlah keterlambatan dan jumlah nasabah bahwa semakin banyak data ditambahkan maka sistem akan makin efisien untuk mengetahui pemetaan wilayah potensial bermasalah

* Prosedur menjalankan aplikasi

Untuk menjalankan program aplikasi Penerepan K-Means untuk pemetaan wilayah potensial bermasalah, yaitu dengan memasukkan url : localhost/k-means

**VII. KESIMPULAN**

Kesimpulan penelitian terkait dengan rumusan masalah dan tujuan penelitian (hipotesis penelitian) berdasarkan hasil penelitian (khususnya hasil pengujian) dan pembahasannya.

1. Penerapan Metode K-Means untuk pemetaan wilayah potensial bermasalah nasabah motor yang diuji kinerjanya dengan metode WhiteBox Testing menghasilkan V(G) = CC = 5, sehinggga dinyatakan bahwa sistem ini telah memenuhi syarat logika pemrograman dan tidak kompleks.
2. Sedangkan dengan pengujian BlackBox menyatakan bahwa sistem telah bebas dari berbagai kesalahan komponen komponennya.

**UCAPAN TERIMA KASIH**

1. Allah SWT atas Ridho dan KaruniaNya.
2. Kedua Orang Tua atas kasih sayang, jerih payah dan doa restunya dalam membesarkan dan mendidik penulis.
3. Dosen dan rekan-rekan mahasiswa fikom 2015.

**REFERENSI**

1. Purwahid Patrik dan Kashadi, Hukum Jaminan, (Semarang: Fakultas Hukum, Universitas Diponegoro, 2008), hal 32
2. Han. J, Kamber M., 2006, Data Mining: Concepts and Techniques, Second Edition. Morgan Kaufman. California.
3. Han. J, Kamber M., 2011, Data Mining: Concepts and Techniques, Waltham: Elsevier Inc.
4. Sulistyono dan Wiwik Sulistiyowati, 2017. Peramalan Produksi dengan Metode Regresi Linear Berganda. Prozima, Vol 1, No. 2, Desember 2017, 82-89 E. ISSN. 2541-5115. Sidoarjo.
5. Farizal, Amar Rachman dan Hadi Al Rasyid, 2014. Model Peramalan Komsumsi Bahan Bakar Jenis Premium Di Indonesia dengan Regresi Linear Berganda. Jurnal Ilmiah Teknik Industri , Vol. 13, No. 2, Desember 2014. Depok.
6. Amrin, 2017. Data Mining Dengan Regresi Linier Berganda Untuk Peramalan Tingkat Inflasi. Jurnal Techno Nusa Mandiri. Vol. XIII, No. 1 Maret 2016.
7. Subekti, 1989, Hukum Acara Perdata, (Bandung : PT. Bina Cipta 1989), hal.128.
8. Subekti, 1996, Hukum Perjanjian, PT. Intermasa, Jakarta. Hal. 130.
9. Prasetyo, E., 2006, Data Mining Konsep dan Aplikasi Menggunakan MATLAB, Andi Yogyakarta. Yogyakarta.
10. Hoffer, Jeffrey A., Ramesh, V., and Topi, Heikki. 2011. Modern Database Management 〖10〗^th Edition. New Jersey: Pearson Education.
11. Witten, I.H. and Frank, E. 2005. Data Mining Practical Machine Learning Tools and Techniques. Second Edition. California: Morgan Kaufman.
12. David, Olson & Yong, Shi. Introduction to Business Data Mining. 2011. International Edition: Mc Graw Hill.
13. Kutner, M.H., C.J. Nachtsheim., dan J. Neter. 2004. Applied Linear Regression Models. 4th ed. New York: McGraw-Hill Companies, Inc.
14. Draper, N. dan Smith, H. 1992. Analisis Regresi Terapan. Edisi Kedua. Terjemahan Oleh Bambang Sumantri. Jakarta: Gramedia Pustaka Utama.
15. Sembiring, R.K. 2003. Analisis Regresi. Edisi Kedua. Bandung: Institut Teknologi Bandung.
16. Gujarati, N.D. 2003. Basic Econometrics. 4th ed. New York: McGraw-Hill Companies, Inc.
17. Widarjono, A. 2007. Ekonometrika: Teori dan Aplikasi untuk Ekonomi dan Bisnis. Edisi Kedua. Yogyakarta: Ekonisia Fakultas Ekonomi Universitas Islam Indonesia.
18. Abdul Munir, Rachmat Aulia dan Yuyun Dwi Lestari, 2015. Analisis Metode Linear Regression untuk Prediksi Penjualan Jamur pada Jamur Karunia Berbasis Web.
19. Sutarbi, Tata. 2013. Analisis Sistem Informasi. Yogyakarta: Andi.
20. Witten, Jeffrey L, et all, Metode Desain & Analisis Sistem, Edisi 6, Edisi International, Mc Graw Hill, Andi, Yogyakarta: 2004.
21. Kadir, Abdul. 2003. Pengenalan Sistem Informasi. Edisi I. Yogyakarta. Andi Yogyakarta.
22. Jogiyanto, HM. 2005. Analisis dan Desain Sistem Informasi : Pendekatan Terstruktur Teori dan Praktek Aplikasi Bisnis. Edisi II. Yogyakarta : Andi Yogyakarta.
23. Hariyanto, Bambang, 2004. Sistem Informasi Basis Data: Pemodelan, Perancangan, dan Terapannya. Informatika, Bandung.
24. Pressman, R.S. 2002. Rekayasa Perangkat Lunak : Pendekatan Praktis (Buku I). Yogyakarta : Andi Yogyakarta.
25. Sugiyono, 2014, Metode Penelitian Kuantitatif, Kualitatif dan R&D, Alfabeta, CV. Bandung.

**BIOGRAFI PENULIS**



***Adrian Ahmad*** lahir di Gorontalo tanggal 08 Juli 1991.

***Haditsah Annur, M.Kom*** Sebagai Pembimbing 1

***Serwin, M.Kom*** Sebagai Pembimbing 2