

IMPLEMENTASI METODE K-NEAREST NEIGHBOR DALAM PERAMALAN PENJUALAN MOBIL BEKAS DI KOTA BATAM

Muhammad Firmansyah Rafsyanzani s¹,
Andi Maslan²

¹Program Studi Teknik Informatika Universitas Putera Batam

²Dosen Program Studi Teknik Informatika Universitas Putera Batam

email: pb140210322@upbatam.ac.id

ABSTRACT

The development of information technology begins to enter the most important part in a field of work, information obtained in a job is general and specific. Computers are one of the results of technological developments that can be used to help all types of work, including transactions in business work, with those in the computer software, a user can perform transaction operations quickly. The use of computers is also needed by a marketing. Car sales of PT. Auto Grace is a company engaged in the sale and purchase of used cars, an area domiciled in the city of Batam. Car sales at PT. Auto Grace is currently still unstable and many cars are not sold and are not maintained due to the lack of consumer interest in buying. So that the company does not experience sales problems, the company leadership must be able to find solutions in a way that is able to predict. So this study aims to predict car sales using the K-NN algorithm. The data set used to predict is taken from sales data from 2017 to 2020. After analysis using the K-NN algorithm, with the help of rapidminer software, the accuracy rate in predicting sales is 63.89% with a precision level of 100%. . The question of take that sale car at year 2021 will experience a decline, due to the outbreak of the Covid-19 virus which has resulted in an economic recession.

Keywords: Data Mining; K-Nearest Neighbor;

PENDAHULUAN

Saat ini banyak sekali masyarakat Indonesia yang memiliki kebutuhan yang berbeda-beda, antara lain anak muda yang menyukai mobil, orang tua yang ingin berkendara bersama keluarga, pengusaha yang ingin mobilnya membawa banyak barang, dan pengusaha yang ingin berkendara dengan nyaman. . Berbagai kebutuhan masyarakat Indonesia sebenarnya dapat memenuhi berbagai kebutuhan dan dengan mudah tersedia di Indonesia. Mobil adalah salah satu jenis kendaraan yang dibutuhkan banyak orang. Umumnya mobil digunakan untuk

dikembangkan sistem informasi berbasis web. menempuh jarak yang cukup jauh. Banyak orang yang bersainglomba membutuhkan sebuah mobil. Khusus di Batam, mobil yang dijual di Batam tidak dibatasi oleh PPN dan PPNBM pemerintah, yang harganya jauh lebih murah dibanding kota lain. Pajak Pertambahan Nilai (PPN) adalah biaya yang dipungut atas pertambahan nilai barang atau jasa yang beredar dari produsen ke konsumen. Namun, pembelian mobil di Batam tanpa PPN tidak bisa sembarangan, misalnya di luar kota Batam. Kalau mau bawa ke luar Batam harus bayar pajak dulu. Banyak

showroom di Batam Mereka membeli mobil dari Singapura yang jenis dan Saat ini banyak sekali masyarakat Indonesia yang memiliki kebutuhan yang berbeda-beda, antara lain anak muda yang menyukai mobil, orang tua yang ingin berkendara bersama keluarga, pengusaha yang ingin mobilnya membawa banyak barang, dan pengusaha yang ingin berkendara dengan nyaman. . Berbagai kebutuhan masyarakat Indonesia sebenarnya dapat memenuhi berbagai kebutuhan dan dengan mudah tersedia di Indonesia. Mobil adalah salah satu jenis kendaraan yang dibutuhkan banyak orang. Umumnya mobil digunakan untuk menempuh jarak yang cukup jauh. Banyak orang yang bersaing.

KAJIAN TEORI

Dalam penelitian ini peneliti melakukan studi pustaka yang berkaitan dengan topik penelitian dan metode penelitian yaitu jurnal dan buku terkait, untuk mencari informasi guna menyusun dan membahas teori-teori terkait...

Data mining adalah analisis pemeriksaan kumpulan data untuk menentukan hubungan yang tidak terduga dan meringkas data dengan cara yang berbeda dari sebelumnya, yang dapat dimengerti dan berguna bagi pemilik data. Penambangan data adalah bidang dari beberapa bidang ilmiah, yang mengintegrasikan pembelajaran mesin, pengenalan pola, informasi statistik, basis data, dan teknologi visualisasi untuk menyelesaikan masalah pengambilan informasi dari basis data besar.

A. Pengelompokan Data Mining
Pengelompokannya adalah sebagai berikut :

1. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, tetapi variabel target mungkin lebih dari

klasifikasi, dan nilainya lebih besar. Buat model menggunakan catatan lengkap yang memberikan nilai variabel target sebagai nilai prediksi. Selain itu, pada pemeriksaan selanjutnya, variabel target diestimasi berdasarkan nilai variabel prediktor. Contohnya adalah memperkirakan kredit kumulatif mahasiswa pascasarjana dengan melihat nilai kredit mahasiswa yang mengikuti program sarjana..

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, dan hanya nilai prediksi yang akan muncul di masa mendatang. Metode dan teknik tertentu yang digunakan dalam klasifikasi dan estimasi juga dapat (jika sesuai) digunakan untuk prediksi. Dalam klasifikasi, ada variabel kategori sasaran. Sebagai contoh, klasifikasi pendapatan dapat dibagi menjadi tiga kategori yaitu berpenghasilan tinggi, berpenghasilan menengah dan berpenghasilan rendah.

4. Pengklusteran

Ini adalah kombinasi dari catatan, pengamatan atau catatan, dan merupakan kelas dari objek serupa. Kluster adalah sekumpulan rekaman yang serupa satu sama lain tetapi tidak serupa dengan rekaman di kluster lain. Perbedaan antara clustering dan klasifikasi adalah tidak ada variabel target dalam clustering. Clustering tidak mencoba untuk mengklasifikasikan, memperkirakan atau memprediksi nilai variabel target. Namun algoritma clustering berusaha membagi semua data menjadi satu kelompok (unity), dimana kemiripan record dalam satu kelompok akan memiliki nilai maksimal, dan kemiripan record pada kelompok lain akan memiliki nilai minimum.

5. Asosiasi

Tugas korelasi dalam data mining adalah menemukan atribut yang hanya muncul satu kali. Dalam dunia bisnis, hal ini sering disebut analisis keranjang belanja. Analisis keranjang pasar adalah metode yang secara khusus menganalisis perilaku suatu kelompok / kelompok konsumen tertentu. Ketika kita tidak mengetahui pola spesifik yang kita cari,

biasanya kita menggunakan analisis keranjang pasar sebagai titik awal untuk mendapatkan pengetahuan tentang perdagangan data. Teknologi analisis keranjang pasar merupakan teknologi yang cocok untuk ilmu data mining. Teknologi ini digunakan untuk merancang strategi penjualan dan pemasaran produk melalui proses pencarian untuk menemukan tautan atau hubungan antar item data dalam database relasional..

A. Metode Data Mining

Analisis asosiasi atau penambangan aturan asosiasi adalah teknik penambangan data yang digunakan untuk menemukan aturan asosiasi antara kombinasi pengukuran minat proyek yang dapat digunakan untuk penambangan data.

1. *Support*

Ukuran yang menunjukkan seberapa besar dominasi suatu item atau kelompok item dalam keseluruhan transaksi.

2. *Confidence*

Menampilkan ukuran hubungan antara dua item dalam suatu kondisi (berdasarkan kondisi tertentu).

B. Algoritma K-Nearest Neighbor

Umumnya, data mining memiliki keahlian untuk klasifikasi. Metode teknis mencakup dua teknik: pembelajaran yang diawasi daripada pembelajaran yang diawasi; pembelajaran yang diawasi bertujuan untuk menemukan pola baru dalam data dengan menghubungkan pola data yang ada dan data baru; pada saat yang sama, dalam pembelajaran tanpa pengawasan, data tidak memiliki pola dan tujuan. Tanpa pengawasan belajar untuk menemukan pola dalam data. Dalam studi penerimaan baru ini, algoritma pembelajaran digunakan diawasi. Metode tetangga K-terdekat bekerja berdasarkan asumsi data akan memiliki kategori atau kategori sama seperti data di sekitarnya

Algoritma K Nearest Neighbor (KNN) adalah metode pengelompokan objek berdasarkan contoh pelatihan terbaru di ruang fitur. Tetangga terdekat K adalah

tipe paling dasar dari pembelajaran berbasis contoh atau pembelajaran malas, dan juga merupakan kelompok pembelajaran berbasis contoh. Dengan menemukan kelompok objek terdekat (serupa) dalam data baru atau data uji dalam data kereta, objek terdekat K dapat diselesaikan. Tetangga terdekat K. Dengan kata lain knn adalah algoritma yang digunakan untuk mengklasifikasikan data berdasarkan data pembelajaran (train data set) yang diperoleh dari k tetangga terdekat. Dimana k adalah jumlah tetangga terdekat. Pembelajaran yang diawasi bertujuan untuk menemukan pola baru dalam data dengan menghubungkan pola data yang ada ke data baru; pada saat yang sama, dalam pembelajaran tanpa pengawasan, data belum menjadi pola, dan target tanpa pengawasan sedang belajar untuk menemukan pola data. Dalam studi pendaftaran ini, algoritma pembelajaran yang diawasi digunakan. Metode K tetangga terdekat mengasumsikan bahwa data dan data sekitarnya termasuk dalam kategori yang sama.

METODE PENELITIAN

1. Waktu dan Lokasi Penelitian
Penelitian akan dilakukan mulai September 2020 hingga Januari 2021 yang berlangsung selama 6 bulan, dan penelitian akan dilakukan di PT Auto Grace..

2. Teknik Pengumpulan Data

Untuk memperoleh data yang diperlukan untuk mendukung penelitian ini, digunakan beberapa teknik pengumpulan data, diantaranya:

Satu jenis. Pengamatan teknis pengumpulan data dengan mengamati langsung objek-objek yang ada di lokasi.

a. Penelitian kepustakaan Dalam penelitian ini, peneliti melakukan penelitian kepustakaan yang berkaitan dengan topik penelitian dan metode penelitian (yaitu jurnal dan buku terkait) untuk mencari informasi guna menyusun dan membahas teori-teori terkait.

b. Wawancara dilakukan dengan pihak-pihak yang terkait dengan topik penelitian yang diusulkan.

3. Desain Penelitian

a. Untuk memperoleh data yang diperlukan untuk mendukung penelitian ini, digunakan beberapa teknik pengumpulan data, diantaranya:

Satu jenis. Pengamatan teknis pengumpulan data dengan mengamati langsung objek-objek yang ada di lokasi.

b. Penelitian kepustakaan Dalam penelitian ini, peneliti melakukan penelitian kepustakaan yang berkaitan dengan topik penelitian dan metode penelitian (yaitu jurnal dan buku terkait) untuk mencari informasi guna menyusun dan membahas teori-teori terkait.

c. Wawancara dilakukan dengan pihak-pihak yang terkait dengan topik penelitian yang diusulkan..

a. Implementasi Dengan Aplikasi d. *Rapidminer*

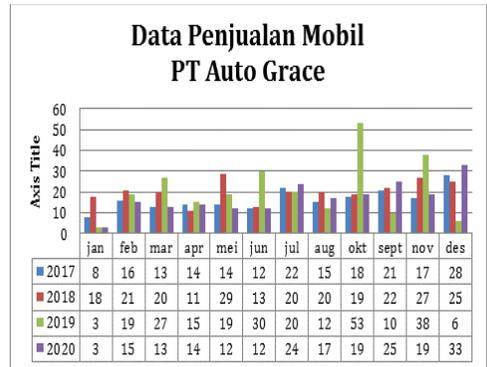
Setelah semua data diolah maka data tersebut diimplementasikan kedalam aplikasi *Rapidminer*. Data yang diolah dengan aplikasi ini akan menjadi pertimbangan dalam perhitungan yang dilakukan secara manual.

HASIL DAN PEMBAHASAN

Data penjualan mobil dari 2017 hingga 2020 akan digunakan untuk pelatihan. Data latih dibagi menjadi dua kelompok, yaitu data masukan dan data sasaran. Data yang diinput adalah data penjualan bulan pertama sampai bulan ke 12, dan data target menggunakan data bulan ke 13. Misal target penjualan Januari 2018, data masukan bulan Januari sampai Desember 2017. Data penjualan bulanan, dan seterusnya, hingga data yang ada dibatasi untuk data penjualan Desember 2020. Data penelitian diambil dari total volume penjualan tahunan dari tahun 2017 sampai tahun 2020. Data tersebut adalah sebagai berikut

	jan	Feb	Mar	apr	mei	jun	jul	aug	okt	sept	nov	des
17	8	16	13	14	14	12	22	15	18	21	17	28
18	18	21	20	11	29	13	20	20	19	22	27	25
19	3	19	27	15	19	30	20	12	53	10	38	6
20	3	15	13	14	12	12	24	17	19	25	19	33

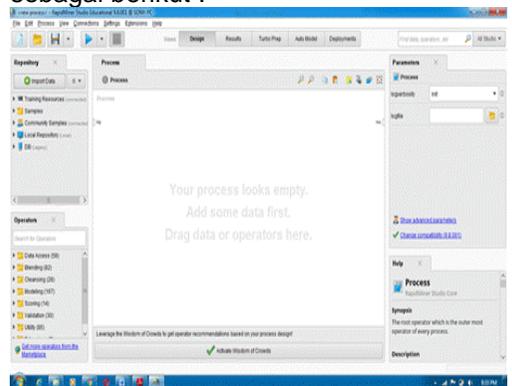
Gambar 1: Tabel Hasil Penjualan



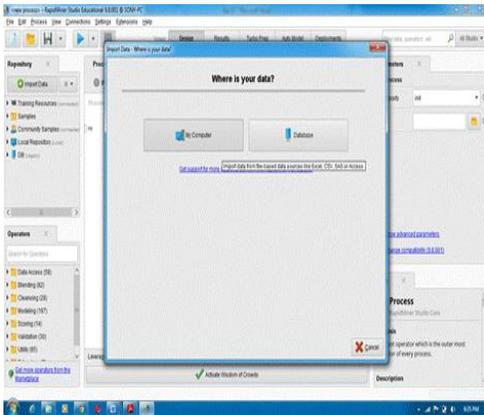
Gambar 2 : Grafik Data Penjualan

1. Model KNN

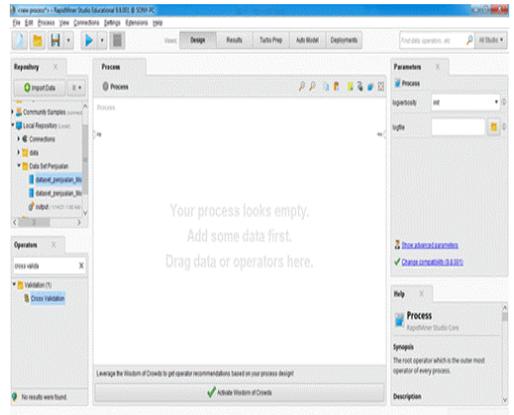
Pada Tahap ini, Pengujian model menggunakan RapidMiner. Cara menggunakan RapidMiner adalah sebagai berikut :



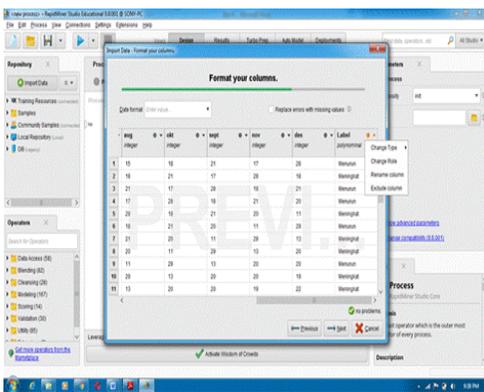
Gambar 3 : Lembar Kerja Rapidminer



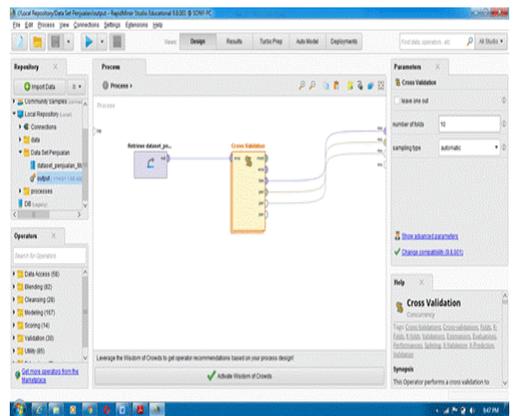
Gambar 4 : Import Dataset



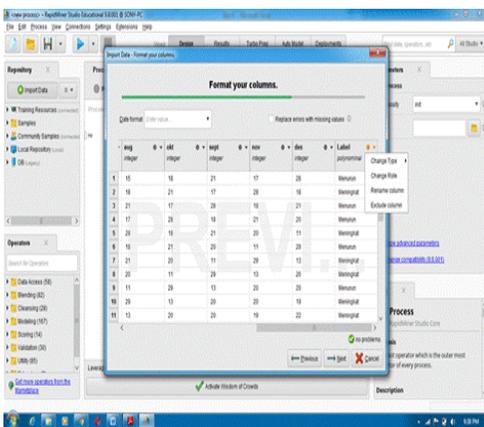
Gambar 7 : Lembar Kerja Proses



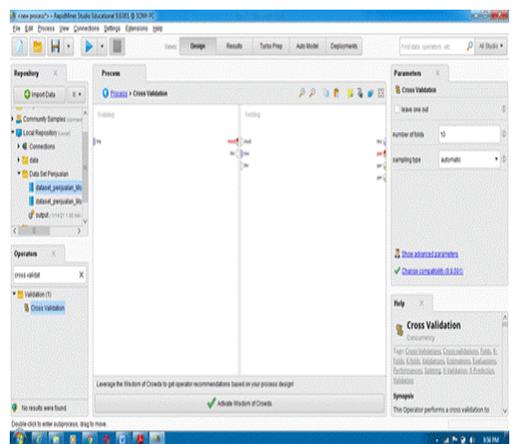
Gambar 5 : Pengaturan Tipe Data Pada Atribut



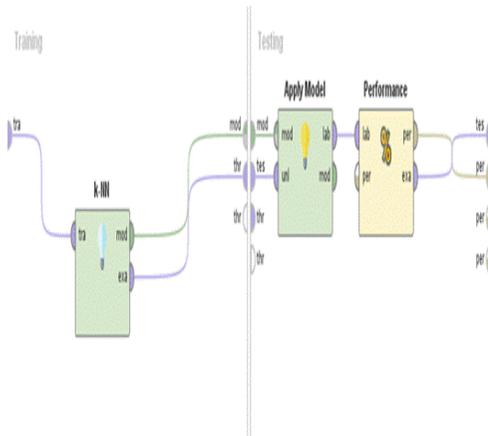
Gambar 8 : Model Cross Validation



Gambar 6 : Hail Import Dataset



Gambar 9 : Lembar Kerja Penempatan Model K-NN



Gambar 10 : Hasil Model K-NN

SIMPULAN

Berdasarkan hasil penelitian berkaitan dengan prediksi penjualan mobil bekas di PT Auto Grace, maka dapat disimpulkan beberapa hal sebagai berikut:

1. Model K-NN dapat digunakan untuk memprediksi penjualan mobil di PT Auto Grace Kota Batam dengan tingkat akurasi mencapai 63.89%.
2. Tingkat Precision dan Recall mencapai 100% artinya model K-NN mampu memprediksi tingkat penjualan mobil dengan benar.
3. Prediksi penjualan mobil pada tahun 2021 berdasarkan penelitian ini akan mengalami penurunan sekitar 36%. Juga dipengaruhi oleh wabah covid-19 yang berdampak pada Indonesia dan seluruh Negara lainnya.
4. Mobil yang laris terjual dapat di analisis dari tahun 2017 sampai dengan 2020 di dominasi oleh mobil merek Toyota.

DAFTAR PUSTAKA

- Abdul Rohman. (2012). *MODEL ALGORITMA K-NEAREST NEIGHBOR (K-NN) UNTUK PREDIKSI KELULUSAN MAHASISWA*. 236 hal.
- Badrul, M. (2016). Algoritma Asosiasi Dengan Algoritma Apriori Untuk Analisa Data Penjualan. *None*, 12(2), 121–129.
- Emha Taufiq Luthfi. (2009). Penerapan Data Mining Algoritma Asosiasi Untuk Meningkatkan Penjualan. *Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III*, 10(1), 1–21.
- Fatoni, C. S., Utami, E., & Wibowo, F. W. (2018). Expert system for diagnosing diphtheria with k-nearest neighbor method. *International Journal Artificial Intelligent and Informatics*, 1(2), 45. <https://doi.org/10.33292/ijarlit.v1i1.4>
- Handoko, K. (2016). Penerapan Data Mining Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode K-Means Clustering (Studi Kasus Di Program Studi Tkj Akademi Komunitas Solok Selatan). *Jurnal Teknologi Dan Sistem Informasi*, 02(03), 31–40. <http://teknosi.fti.unand.id/index.php/teknosi/article/view/70>
- Ilmu, F., & Universitas, K. (2018). *Perbandingan Metode K-Nearest Neighbor dan Metode Naive Bayes untuk Klasifikasi Komentar Produk Online Shop Berdasarkan Ulasan Pembeli*.
- KumarShrivastava, A., & Kumar Dewangan, A. (2014). An Ensemble Model for Classification of Attacks with Feature Selection based on KDD99 and NSL-KDD Data Set. *International Journal of Computer Applications*, 99(15), 8–13. <https://doi.org/10.5120/17447-5392>
- Kurniawan, D., & Saputra, A. (2019). Penerapan K-Nearest Neighbour dalam Penerimaan Peserta Didik dengan Sistem Zonasi. *Jurnal*

- Sistem Informasi Bisnis*, 9(2), 212.
<https://doi.org/10.21456/vol9iss2pp212-219>
- Lubis, A. H. (2016). Model Segmentasi Pelanggan Dengan Kernel K-Means Clustering Berbasis Customer Relationship Management. *Jurnal & Penelitian Teknik Informatika*, 1, 36–41.
- Maslan, A., Mohamad, K. M. Bin, & Mohd Foozy, F. B. (2020). Feature selection for DDoS detection using classification machine learning techniques. *IAES International Journal of Artificial Intelligence*, 9(1), 137–145.
<https://doi.org/10.11591/ijai.v9.i1.pp137-145>
- Mustakim, & Oktaviani F, G. (2016). *Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa*. 13(2), 195–202.
- Sambanyu, A. P. (2016). *Padat dan Potensi Sebagai Co - Firing Agent*. 5(2), 70–73.
- Saxena, K., Khan, Z., & Singh, S. (2014). Diagnosis of Diabetes Mellitus using K Nearest Neighbor Algorithm. *International Journal of Computer Science Trends and Technology (IJCST)*, 2(4), 36–43.
- Sugesti, A., Mukid, M. A., & Tarno, T. (2019). Perbandingan Kinerja Mutual K-Nearest Neighbor (Mknn) Dan K-Nearest Neighbor (Knn) Dalam Analisis Klasifikasi Kelayakan Kredit. *Jurnal Gaussian*, 8(3), 366–376.
<https://doi.org/10.14710/j.gauss.v8i3.26681>
- Thi Bi Dan, T., Widya Sihwi, S., & Anggrainingsih, R. (2016). Implementasi Iterative Dichotomiser 3 Pada Data Kelulusan Mahasiswa S1 Di Universitas Sebelas Maret. *Jurnal Teknologi & Informasi ITSmart*, 4(2), 84.
<https://doi.org/10.20961/its.v4i2.1770>
- Tiaratuti, A. D. (2014). *Peramalan Penjualan Mobil Pada Pt Bengawan*. 8.
- Wijayanti, A. W. (2017). Analisis Hasil Implementasi Data Mining Menggunakan Algoritma Apriori pada Apotek. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 3(1), 60.
<https://doi.org/10.26418/jp.v3i1.19534>
- Yanto, R., & Khoiriah, R. (2015). Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat. *Creative Information Technology Journal*, 2(2), 102.
<https://doi.org/10.24076/citec.2015v2i2.41>

BIODATA PENULIS

 A portrait of a young man with short dark hair, wearing a blue and yellow striped polo shirt, against a light blue background.	<p>Biodata' Penulis pertama, M.Firmansyah Rafsyanzani S, merupakan mahasiswa Prodi Teknik Informatika Universitas Putera Batam.</p>
 A portrait of a man with short dark hair, wearing a dark blue suit jacket, a white shirt, and a striped tie, against a red background.	<p>Biodata' Penulis kedua, Andi Maslan, merupakan Dosen Prodi Teknik Informatika Universitas Putera Batam.</p>