# Mining Data Kependudukan untuk Validasi Identitas Calon Penerima Bantuan Jaring Pengaman Sosial di Kabupaten Gresik

[1]Khurin In Noviarani, [2]Umi Chotijah

Teknik Informatika, Universitas Muhammadiyah Gresik
Jl. Sumatera No.101, GKB, Randuagung, Kec. Kebomas, Kabupaten Gresik, Jawa Timur, Indonesia

Email: [1]khurin.noviarani@gmail.com, [2]umi.chotijah@umg.ac.id

**Tersedia Online di**
http://www.jurnal.unublitar.ac.id/index.php/briliant

**Sejarah Artikel**
Diterima pada Desember 2020
Disetuji pada Februari 2021
Dipublikasikan pada Mei 2021
Hal. 416-428

**Kata Kunci:**
Data Mining; Klasifikasi; *Decision Tree*

**DOI:**
http://dx.doi.org/10.28926/briliant.v3i4.598

**Abstrak**: Tulisan ini memuat gambaran penggunaan *data mining* model klasifikasi dengan teknik *decision tree* untuk membantu proses *validasi* identitas calon penerima bantuan secara cepat dan akurat, sehingga dapat membantu memudahkan pemangku kepentingan dalam hal ini Pemerintah Kabupaten Gresik untuk mengambil keputusan dalam melakukan penyaluran subsidi Jaring Pengaman Sosial (JPS) dampak pandemi COVID-19 secara merata dan tepat sasaran sesuai dengan kebijakan yang dibuat. Proses *validasi* dimulai dengan memeriksa data apakah Nomor Induk Kependudukan (NIK) kosong atau tidak, dilanjutkan setiap cabangnya menunjukkan kemungkinan skenario dari keputusan yang diambil dan hasil akhirnya adalah data tersebut statusnya *valid* atau statusnya tidak *valid*. Penggunaan *data mining* model klasifikasi dengan teknik *decision tree* dalam tulisan ini membuktikan bahwa penggunaan tehnik tersebut dapat menghemat 50% usaha / *effort* yang dilakukan dan mempercepat proses sebesar 31 Jam atau 89,84702093% dari proses semula yang membutuhkan 34,5 jam kerja. Otomatisasi system menggunakan *data mining* model klasifikasi dengan teknik *decicion tree* sangat membantu mempermudah, mengefisienkan waktu dan meminimalkan *human error* dalam pekerjaan validasi identitas calon penerima bantuan JPS di Kabupaten Gresik.

## PENDAHULUAN

Pandemi COVID-19 di tahun 2020 memberikan dampak pada berbagai sektor terutama ekonomi, yang membuat berkurang hingga kehilangan mata pencaharian. Untuk menjaga stabilitas ekonomi dan menjaga kelangsungan kehidupan warga negaranya pemerintah mengambil kebijakan dengan memberikan subsidi langsung kepada warganya.

Untuk memastikan pemerataan calon penerima manfaat subsidi Jaring Pengaman Sosial (JPS) dalam kebijakannya pemerintah khususnya Pemerintah Kabupaten Gresik memberikan beberapa prasyarat yang antara lain: Bukan Aparatur Pemerintah (ASN/TNI/POLRI). Prioritas utama adalah lansia, janda, dan masyarakat rentan miskin. Bukan penerima aktif bantuan dari pemerintah. Bantuan bukan untuk individu namun untuk keluarga (Umum et al., 2016).

Penyaringan calon penerima bantuan yang diajukan perlu dilakukan agar prasyarat dalam kebijakan terpenuhi. Dalam proses penyaringan calon penerima bantuan kecepatan dalam pengambilan keputusan tepat dan akurat sangatlah

penting, mengingat penyebaran Pandemi COVID dan dampaknya yang sangat cepat.

Pada awal pelaksanaan kebijakan di Kabupaten Gresik, penyaringan dilakukan secara manual dengan memanfaatkan aplikasi pengolah angka (*spreadsheet*) menggunakan fungsi bawaan seperti *lookup* dan *if*. Namun dalam penggunaan aplikasi pengolah angka untuk kegiatan penyaringan memakan banyak waktu dan tingkat kebenaran dan keakuratan hasil penyaringan data sangat bergantung pada mental dan ketelitian operator / *Person In Charge (PIC)* penyaring data.

Untuk meningkatkan akurasi dan kecepatan proses penyaringan data calon penerima subsidi, diperlukan *upgrade* alat yang semula hanya menggunakan aplikasi pengolah angka ke aplikasi custom yang dapat melakukan penggalian data identitas, data sosial dan data penerima subsidi, sekaligus menentukan sah tidaknya penerima bantuan yang diajukan.

*Data mining* bisa digunakan untuk memperbaiki proses penyaringan data yang dilakukan secara manual sehingga bisa menangani: (1) Jumlah data yang besar. (2) Dimensi data yang tinggi. (3) Data yang heterogen dan beda memiliki sifat yang berbeda (*Decision Tree Algorithm Examples In Data Mining*, n.d.; Larose, 2005; Turban, 2005).

Tulisan ini akan memberikan gambaran penggunaan *data mining model klasifikasi* dengan tehnik *Decision tree* yang merupakan salah satu teknik dalam *data mining* untuk mengelompokkan data berdasarkan keterikatan data terhadap data sampel, dalam aplikasi custom untuk validasi identitas penerima bantuan secara cepat dan akurat, melalui otomatisasi validasi data identitas menggunakan data mining model klasifikasi dengan teknik decision tree, sebagai alat bantu pengambilan keputusan dalam penyaluran subsidi Jaring Pengaman Sosial (JPS) dampak pandemi COVID, agar subsidi tersebut dapat tersalurkan dengan merata dan tepat sasaran sesuai dengan kebijakan yang dibuat.

Data mining model klasifikasi dengan tehnik Decision tree banyak digunakan dalam proses analisa data baik untuk proyeksi maupun pengambilan keputusan dalam berbagai bidang seperti dicontohkan dalam bidang pendidikan untuk klasifikasi penjurusan atau prediksi kelulusan dalam bidang industri untuk penjadwalan shift kerja atau proyeksi keuntungan, sedang di bidang kesehatan tehnik ini banyak digunakan untuk mendeteksi wabah dan pola penyakit melalui analisa rekam medis atau pemilihan pola diet untuk pasien.

Namun sayangnya tehnik ini jarang digunakan dalam bidang permerintahan. Penulis berharap hasil dari penelitian ini dapat memberikan gagasan baru melalui contoh kecil yang dilakukan penulis dalam pemanfaatan data kependudukan dan konsep data mining di bidang pemerintahan sebagai sistem penunjang pembuatan kebijakan.

## METODE

Penelitian ini merupakan jenis penelitian analisis deskriptif kualitatif dimana dilakukan pengumpulan data, analisa dan kesimpulan. Analisa akan didefinisikan menjadi: Analisa kebutuhan, Analisa proses dan Analisa output.

Metode pengumpulan data yang digunakan meliputi: *Review* Dokumen, *Observasi* dan *Interview*. Dalam proses review dokumen, dilakukan penggalian informasi yang tersaji dalam teori dan latar belakang yang telah dikumpulkan dari berbagai sumber antara lain jurnal akademik, dokumen pemerintah, user manual dan kertas kerja atau catatan. Dalam proses ini dilakukan pula analisa template dokumen aplikasi pengolah angka yang sebelumnya digunakan petugas data screening atau *PIC (Person In Charge)* untuk melakukan penyaringan secara manual. Sedangkan observasi dilakukan untuk mencari data bagaimana proses validasi data calon penerima bantuan dilakukan. Metode observasi memungkinkan untuk mengamati fakta dari sudut pandang *subject* sehingga bisa diperoleh data yang *objektif*. Kemudian *interview* yang merupakan metode pengumpulan data dengan jalan tanya jawab dengan *PIC* atau pihak-pihak yang bersangkutan untuk memperoleh informasi yang dibutuhkan.

Analisa Kebutuhan mendefinisikan komponen yang diperlukan untuk menganalisa kebutuhan dari aplikasi antara lain basis data dan sumber data, kemudian masukan (input). Dalam tahap ini didefinisikan apa saja basis data yang dibutuhkan untuk *upgrade* alat yang semula hanya menggunakan aplikasi pengolah angka ke aplikasi custom yang dapat melakukan penggalian data identitas, data sosial dan data penerima subsidi, sekaligus menentukan sah tidaknya penerima bantuan yang diajukan sekaligus sumber datanya dari mana. Kemudian didefinisikan juga masukan apa yang dibutuhkan agar aplikasi custom dapat bekerja sebagaimana mestinya. Analisa Proses mendefinisikan bagaimana aplikasi custom ini dapat bekerja membantu *PIC* melakukan penggalian data identitas, data sosial dan data penerima subsidi, sekaligus menentukan sah tidaknya penerima bantuan. Analisa Output menggambarkan keluaran dari aplikasi custom yang dapat dipahami oleh *PIC*.

Teori yang menjadi landasan dalam tulisan antara lain: Teori *Data Mining* dan Teori *Decision Tree*.

*Data Mining* bertujuan untuk memperbaiki teknik tradisional sehingga bisa menangani: Jumlah data yang sangat besar, dimensi data yang tinggi, data yang heterogen dan memiliki sifat yang berbeda (Novianti & Santosa, 2016). Dalam penelitian ini penulis memilih menggunakan tehnik klasifikasi dari berbagai tehnik *data mining* yang ada, tehnik klasifikasi ini merupakan tehnik *data mining* untuk membangun model yang dapat mengklasifikasikan suatu objek berdasar atribut-atributnya. Dalam kasus screening data calon penerima subsidi Jaring Pengaman Sosial (JPS) dampak pandemi COVID pemilihan data mining dengan tehnik klasifikasi merupakan pilihan yang tepat, karena proses screening yang dilakukan PIC merupakan proses pemilihan dan klasifikasi data.

*Decision Tree* adalah pohon dimana setiap cabangnya menunjukkan pilihan diantara sejumlah alternatif pilihan yang ada, dan setiap daunnya menunjukkan keputusan yang dipilih (Ariestya et al., 2016; Durrotul, 2017; Kasih, 2019; Rismayanti, 2018; Sutoyo, 2018). *Decision tree* biasa digunakan untuk mendapatkan informasi untuk tujuan pengambilan sebuah keputusan. *Decision tree* dimulai dengan sebuah *root node* (titik awal) yang dipakai oleh *user* untuk mengambil tindakan. Dari node root ini, *user* memecahnya sesuai dengan *algoritma decision tree*. Hasil akhirnya adalah sebuah *decision tree* dengan setiap cabangnya menunjukkan kemungkinan skenario dari keputusan yang diambil serta hasilnya.

## HASIL DAN PEMBAHASAN

Berdasarkan hasil pengumpulan data yang berupa *Review* Dokumen, *Observasi* dan *Interview* yang telah dipelajari untuk mengenali apa itu proses Penyaringan calon penerima subsidi Jaring Pengaman Sosial (JPS) dampak pandemi COVID? bagaimana memanfaatkan *data mining* model klasifikasi dengan teknik *Decision tree* dalam proses ini? apa saja yang dibutuhkan untuk memanfaatkan *data mining* model klasifikasi dengan teknik *Decision tree* dalam proses ini? disimpulkan bahwa pemanfaatan terbaik adalah dengan membuat aplikasi berbasis website semudah mungkin untuk digunakan.

### 1. Analisa Kebutuhan

Dalam membuat aplikasi berbasis website memanfaatkan *data mining* model klasifikasi dengan teknik *Decision tree* ini sebelumnya dilakukan analisa kebutuhan untuk mengenali Sistem yang dibutuhkan, Basis Data dan Sumber data yang dimiliki, dan kebutuhan masukan / *input* tambahan dari pengguna, yang secara detail akan dibahas pada poin - poin dibawah ini:

#### a. Kebutuhan Sistem

Agar dapat digunakan secara efisien, system minimal membutuhkan Hardware Server dengan Quad Core Processor dengan 8 Threads, 6MB Cache dan 2.2 Ghz clock speed. Dengan RAM 8GB dengan BUS Rate 1333 MT/s, Bandwidth 10666 MB/s, BUS Clock Rate 666Mhz, dan Cycle time tidak lebih dari 6 ns. Untuk Storage menggunakan SAS Hard Drive dengan kapasitas paling Tidak 500 GB dengan bandwith 6GB/s. Sedangkan untuk Software Operating System Linux Debian 9 64bit, yang terinstal web server Apache, CURL, MariaDB, dan PHP 7, php7-mysql, libapache2-mod-php7, php7-common, libphp7-embed, php7-json, php7-CLI, php7-opcache, php7-readline, php7-mbstring, php-curl, php-bcmath, php-bz2, php-enchant, php7-gmp, php7.3-xml, php-xml-htmlsax3.

#### b. Basis Data dan Sumber data

Beberapa data yang dibutuhkan dalam proses data mining antara lain sebagaimana pada tabel 1 basis data dan sumber data di bawah ini:

Tabel 1. Basis Data dan Sumber Data

| Basis Data | Sumber Data | Keterangan |
|---|---|---|
| Data Kependudukan | Data kependudukan diperoleh dari Server Sistem Informasi Kependudukan milik Dinas Kependudukan dan Pencatatan Sipil Kabupaten Gresik melalui *API* yang disediakan berdasarkan Nota Kesepahaman dan Perjanjian Penggunaan antara Dinas Kependudukan dan Pencatatan Sipil dan Badan Perencanaan Pembangunan, Penelitian dan Pengembangan Daerah Kabupaten Gresik | Data Kependudukan merupakan data inti yang digunakan dalam proses validasi, dengan memanfaatkan elemen data dari response *API* yang antara lain : (1). Response Status (NIK ditemukan/Tidak). (2). Nomor KK. (3). Nama. (4). Alamat. (5). Tanggal Lahir. (6). Jenis Kelamin. (7).Pekerjaan. (8).Status Perkawinan. |

| | | |
|---|---|---|
| Data Sosial / BDT (Basis Data Terpadu) | Data Sosial diperoleh dari Server Database Aplikasi GRANDE yang merupakan kompilasi export data aplikasi SIKS-NG Kementrian Sosial yang dikelola oleh Dinas Sosial Kabupaten Gresik | Data ini dipergunakan untuk melakukan validasi penerima subsidi sosial lain yang bersumber dari kementrian sosial seperti PKH, BSP, dan BST |
| Data Penerima Bantuan JPS | Data ini merupakan import data Aplikasi Pengolah Angka yang diperoleh dari desa pengusul dan penerima subsidi JPS yang telah disahkan | |
| Data *Cache* Kependudukan | | Untuk mempersingkat pencarian data kependudukan, diperlukan data cache yang akan mengurangi request NIK berulang ke server Sistem Informasi Kependudukan. data cache disimpan selama sesi user aktif dan akan dihapus setelah sesi user tidak aktif, mengingat regulasi tidak memperbolehkan mirroring dan menyimpan data kependudukan dari SIAK |
| Data Janda | Data janda merupakan import data Aplikasi pengolah angka yang berasal dari desa/ kelurahan dan Bagian Kesejahteraan Rakyat Sekretariat Daerah Kabupaten Gresik | |
| Data Pekerjaan *Blacklist* | Import data Aplikasi Pengolah Angka pekerjaan yang dilarang untuk menerima bantuan JPS | |

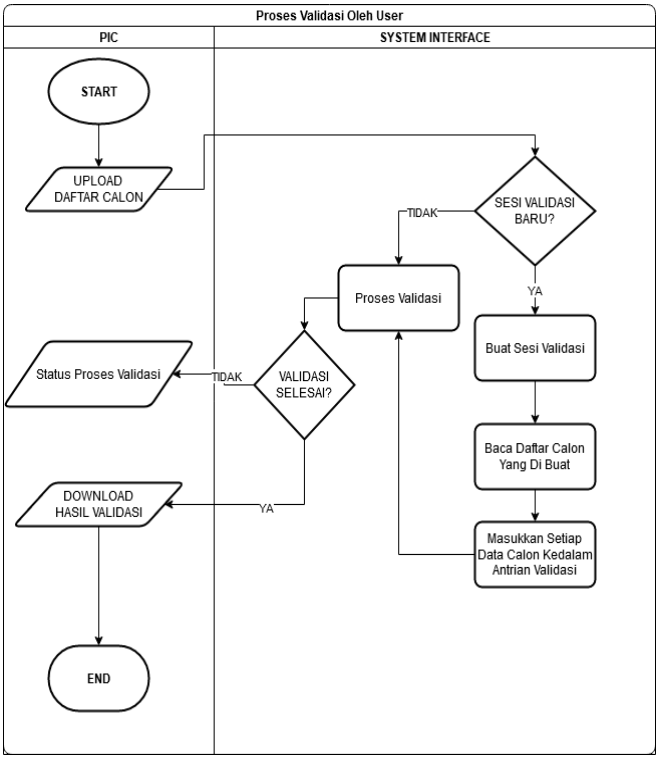
**c. Masukan (*Input*)**

Untuk memudahkan pengguna (PIC) *input* applikasi dibuat semudah dan sesedikit mungkin yang hanya dilakukan menggunakan *upload* file *template* yang sudah ada dan dua *combobox* yang berisi nama kecamatan dan nama desa/ kelurahan pengusul untuk menjaga konsistensi data/ kelurahan pengusul, dan satu tombol upload untuk mengirim data dan menginstruksikan sistem untuk melakukan *extract* elemen data berikut yang selanjutnya akan diproses oleh sistem.

1. *Main Input*
   a. Nomor Induk Kependudukan
   b. Nomor Kartu Keluarga
2. *Complementary Input*
   a. Nama Calon Penerima
   b. Pekerjaan Calon Penerima
   c. Alamat Domisili Calon Penerima

## 2. Analisa Proses

Secara sederhana proses penggunaan aplikasi berbasis website memanfaatkan *data mining* model klasifikasi dengan teknik *Decision tree* ini, dapat digambarkan melalui diagram alur / *flowchart* pada gambar 1 dibawah ini, sedang bagaimana cara aplikasi melakukan mining secara detail akan di jelaskan pada point (a) peta komunikasi dan poin (b) aliran data sistem.




Gambar 1. *Flowchart* Validasi

pada *flowchart* proses validasi diatas dapat dilihat bahwa proses yang dilakukan PIC (manual oleh manusia) hanya melakukan upload daftar calon penerima, cek status proses validasi, dan download hasil validasi. Selebihnya proses dilakukan secara otomatis oleh sistem.

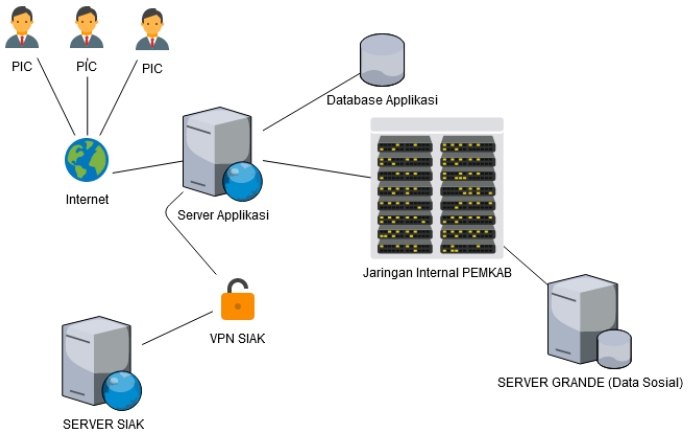
Dalam proses validasi data calon penerima subsidi Jaring Pengaman Sosial (JPS) dampak pandemi COVID sendiri secara umum dan sederhana dijabarkan melalui pseudocode berikut:

1. PIC upload file excel data mentah dengan template yang telah ditentukan.
2. System mulai membaca Kolom Nomor NIK dan KK.
3. Check apabila sesi PIC sudah pernah ada, apabila ada system akan mengambil data cache kependudukan
4. Jika tidak maka System akan mengambil data kependudukan dari SIAK menggunakan NIK
5. Jika data ditemukan maka sistem akan memberikan validasi kepada NIK dan KK dan memberikan data NIK, KK, Nama, Alamat, Tanggal Lahir, Jenis Kelamin, Pekerjaan dari server SIAK.

6. Selanjutnya dengan menggunakan NIK dan KK, system akan mengambil data PKH, BSP, BST dari server Grande
7. Jika data ditemukan maka sistem akan memberikan data ID PKH, ID BSP, ID BST dari server Grande.
8. Kemudian sistem akan mengambil data JPS, DD, Janda Miskin, Pekerjaan menggunakan NIK, KK dan Pekerjaan.
9. Jika data ditemukan maka sistem akan memberikan Status JPS, Status DD, Status Janda Miskin, Validasi Pekerjaan.
10. PIC mendownload file excel hasil validasi.

**a. Peta Komunikasi**

Sistem validasi ini menggunakan interface web yang dapat diakses melalui internet oleh PIC validator untuk melakukan upload data mentah dan download hasil validasi. Proses validasi dilakukan melalui *mining* data secara otomatis ke 3 sumber data yakni Server SIAK melalui jalur VPN SIAK, Server Grande melalui jalur LAN internal data center pemerintah kabupaten Gresik, dan Query langsung ke dalam databaseyang terdapat pada server aplikasi sebagaimana tergambar dalam gambar 2 peta komunikasi berikut:

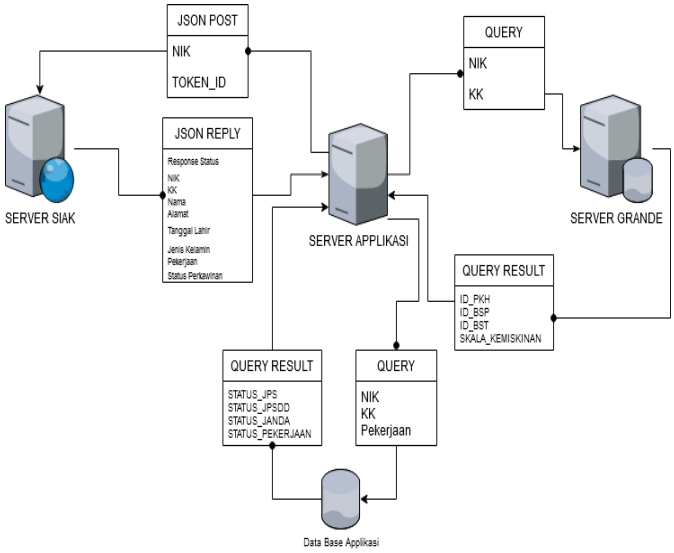



Gambar 2. Peta Komunikasi

**b. Aliran Data Sistem**

Aliran data sebagaimana tergambar dalam gambar 3 menggambarkan bagaimana proses transaksi data yang digunakan oleh server aplikasi sebagai bahan proses validasi menggunakan decision tree. pada gambar dapat dilihat terdapat dua metode transaksi data dilakukan yakni, (1). metode cURL POST dengan request data berformat JSON yang berisi NIK dan TOKEN_ID ke Server SIAK untuk mendapatkan JSON Response yang memiliki elemen data (a). Response Status (NIK ditemukan/Tidak). (b). Nomor KK. (c). Nama. (d). Alamat. (e). Tanggal Lahir. (f). Jenis Kelamin. (g).Pekerjaan. (h).Status Perkawinan. (2). Metode kedua adalah query langsung ke database menggunakan keyword NIK, KK, dan Pekerjaan ke dalam database internal yang ada pada server aplikasi untuk mendapatkan status Janda, Pekerjaan Blacklist, penerima bantuan JPS dan JPSDD , selain itu query langsung juga dilakukan menggunakan keyword NIK dan KK ke Server Database GRANDE

untuk mendapatkan data Sosial yang berupa ID PKH, BSP, BST dan Nilai Skala Kemiskinan.




Gambar 3. Aliran Data Sistem

## 3. Analisa Output

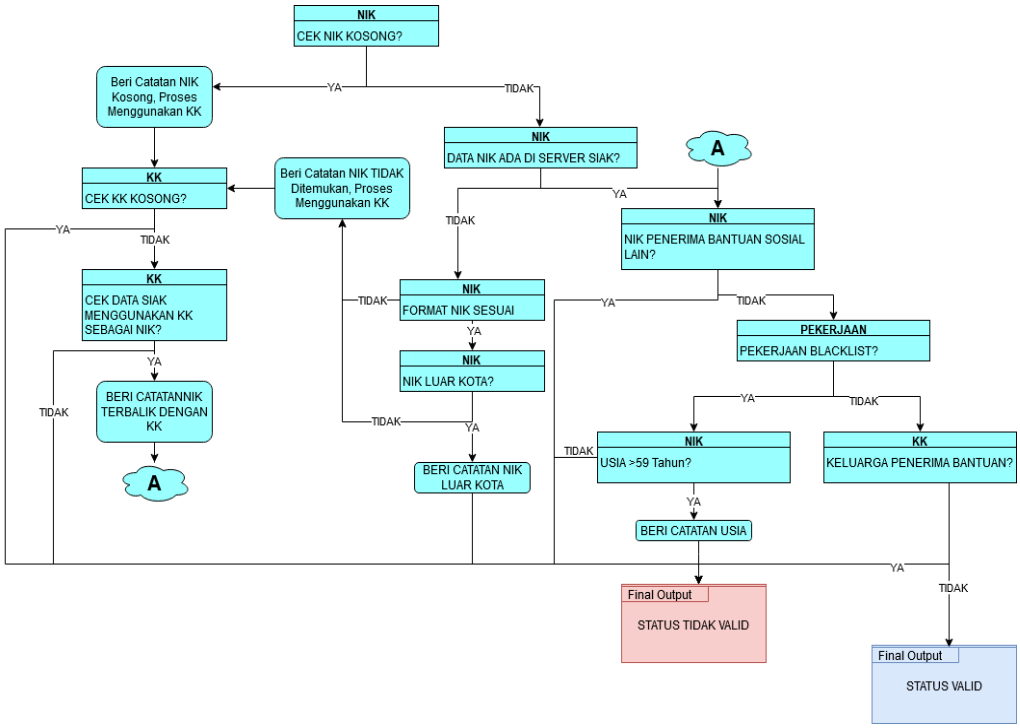
Hasil akhir dari aplikasi ini merupakan file excel template yang di tambahkan kolom dengan isi sebagaimana tabel 2 analisa output berikut:

Tabel 2. Analisa Output

| Kolom Tambahan | Isi dari Kolom Tambahan |
|---|---|
| Hasil Validasi | Valid / Tidak Valid |
| PKH | Ya / Tidak |
| BSP | Ya / Tidak |
| BST | Ya / Tidak |
| ID DTKS | Nomor DTKS |
| ID PKH | Nomor PKH |
| Usia | Umur Calon Penerima |
| Identitas | Pekerjaan, NIK, KK, Nama dan Domisili berdasarkan data SIAK |
| Catatan | (Hanya akan muncul sesuai kondisi yang ada dan bisa tertulis lebih dari 1 catatan)<br>1. NIK tidak valid<br>2. nik tidak ditemukan<br>3. nik luar kota<br>4. nik terbalik dengan kk<br>5. nik keluarga penerima bantuan sosial<br>6. nik keluarga penerima bantuan blt-dd |

| | |
|---|---|
| | 7. nik keluarga penerima jps tahap X<br>8. nama tidak sesuai siak<br>9. alamat tidak sesuai siak<br>10. pekerjaan tidak sesuai siak<br>11. nik tidak sesuai siak<br>12. kk tidak sesuai siak<br>13. pekerjaan siak terdaftar dalam pekerjaan blacklist.<br>14. nik terdaftar dalam keluarga dengan pekerjaan blacklist |

Untuk mencapai hasil akhir yang berupa file excel yang di download, dapat digambarkan melalui *decision tree diagram* yang di ilustrasikan pada gambar 4 berikut




Gambar 4. *Decision Tree*

Pada gambar 4 dapat dilihat bahwa proses validasi dimulai dengan memeriksa data apakah NIKnya kosong atau tidak, dilanjutkan setiap cabangnya menunjukkan kemungkinan skenario dari keputusan yang diambil dan hasil akhirnya adalah data tersebut statusnya valid atau statusnya tidak valid.

## 4. Hasil Penggunaan Aplikasi

Dapat dilihat dari tabel 3 di bahwa penggunaan data *mining* dan otomatisasi dapat menyederhanakan, mempercepat dan memudahkan proses validasi data calon penerima, yang seharusnya dilakukan dalam 6 proses dilakukan dalam 3 proses saja yang berarti menghemat 50% usaha / *effort* yang dilakukan oleh PIC dan mempercepat proses sebesar 31 Jam atau 89,84702093% dari proses semula yang membutuhkan 34,5 jam kerja.

Tabel 3. Hasil

| No | Aksi | Manual | Data Mining | Keterangan |
|---|---|---|---|---|
| 1 | Mengirimkan permintaan data ke OPD Terkait dan mengumpulkan data dari OPD Terkait | Dilakukan | Tidak Dilakukan | Merupakan proses yang harus dikerjakan, namun pada proses otomatisasi data mining hanya perlu 1 payung hukum / MOU, jalur akses yang disediakan opd, dan laporan rekap data yang digunakan. |
| 2 | Mengirimkan data dari OPD ke Desa/ Kelurahan untuk dilakukan validasi domisili, identifikasi penerima Bantuan DD dan mengumpulkan data dari Desa/ Kelurahan | Dilakukan | Dilakukan | Merupakan proses yang harus dikerjakan |
| 3 | Waktu yang dibutuhkan untuk mengolah data dari desa/ kelurahan ke dalam format template JPS | 1 Jam / Batch | 1 Jam / Batch | |
| 4 | Waktu yang dibutuhkan untuk upload file template ke dalam aplikasi data mining | - | 5 Detik / Batch | |
| 5 | Waktu yang dibutuhkan untuk mempersiapkan rumus if dan lookup pada data NIK dan KK untuk validasi identitas calon penerima bantuan | 30 Menit / Batch Data | - | Dilakukan otomatis pada saat data mining mengirimkan permintaan ke server Dispendukcapil |
| 6 | Waktu yang dibutuhkan untuk verifikasi dan validasi identitas calon penerima bantuan | 3 Hari/Batch Data | 1 detik / NIK | Jika dirata-rata dalam batch data terdapat 3000 NIK maka data mining dapat memproses per batch data yang dikirimkan ke dispendukcapil dalam 1 jam (1 jam = 3600 detik) |

| | | | | |
|---|---|---|---|---|
| 7 | Waktu yang dibutuhkan untuk verifikasi dan validasi penerima bantuan sosial | 2 jam / Jenis Bantuan Sosial<br><br>Ada 3 jenis bantuan (PKH, BSP, BST)<br><br>Jadi 6 Jam / Batch Data | 1 detik / NIK | Jika dirata-rata dalam batch data terdapat 3000 NIK maka data mining dapat memproses per batch data yang dikirimkan ke Server Grande dalam 1 jam (1 jam = 3600 detik) |
| 8 | Waktu yang dibutuhkan untuk validasi data | 3 Jam / Batch Data<br><br>Cek status JPS, DD, Janda Miskin, Pekerjaan | 1 detik / NIK | Jika dirata-rata dalam batch data terdapat 3000 NIK maka data mining dapat memproses per batch data yang diproses dalam 1 jam (1 jam = 3600 detik) |
| 9 | Waktu yang dibutuhkan untuk download hasil validasi | - | 5 Detik / Batch | |
| 10 | Jumlah proses yang dilakukan oleh PIC | 6 proses | 3 proses | Jumlah aksi (aksi nomer 1-9) yang dilaksanakan oleh PIC, dalam satuan proses (aksi = proses). |
| 11 | Jumlah Waktu Proses Penyelesaian Skrening | 34,5 Jam | 3,5027777 78 Jam | Total waktu kerja dalam satuan jam untuk aksi nomer nomer 3 – 9 per batch data (3000 NIK) |

## KESIMPULAN

Penggunaan data *mining* dan otomatisasi dapat menyederhanakan, mempercepat dan memudahkan proses validasi data calon penerima, dari berbagai langkah manual validasi melalui penyandingan data fungsi *lookup* dan *if* aplikasi pengolah angka, menjadi hanya 3 proses yang dilakukan oleh pic yaitu : *(1). Upload, (2). Cek Status, (3) Download* seperti yang telah digambarkan dalam *flowchart* gambaran proses validasi di bab pembahasan poin ke 4 dan mempercepat proses yang semula 34,5 jam kerja menjadi 3,5 jam.

Dalam proses otomatisasi validasi, proses utama yang menjadi inti merupakan proses *mining* data informasi kependudukan yang berada di server SIAK Dinas Kependudukan dan Pencatatan Sipil Kabupaten Gresik karena regulasi mensyaratkan bahwa penerima merupakan warga negara yang sah (memiliki NIK yang terdaftar aktif di dalam sistem informasi kependudukan).

Meski kemudahan, kesederhanaan dan kecepatan yang ditawarkan dalam proses validasi dan otomatisasi ini, terdapat kendala non-teknis yang masih ditemukan antara lain: (1) Penulisan NIK dan No KK yang tertukar. (2) Penulisan NIK yang kurang valid (kurang dari 16 digit, atau digit terakhir terkena pembulatan

dari aplikasi pengolah angka). (3) Informasi kependudukan yang tidak sesuai dengan realita (tidak sama dengan yang diajukan oleh perangkat desa, terutama pada data pekerjaan). Data mining khususnya data mining kependudukan di server SIAK memiliki banyak peluang ruang pemanfaatan dalam dunia pemerintahan seperti clustering penduduk berdasarkan usia dan lapangan kerja, estimasi laju penduduk, validasi usulan rencana pembangunan daerah, hingga pencegahan kriminal.

## SARAN

Mengingat sebagian kendala yang dialami dalam proses otomatisasi validasi melalui mining data kependudukan ini adalah kendala non-teknis seperti (1) Penulisan NIK dan No KK yang tertukar. (2) Penulisan NIK yang kurang valid (kurang dari 16 digit, atau digit terakhir terkena pembulatan dari aplikasi pengolah angka). (3) Informasi kependudukan yang tidak sesuai dengan realita (tidak sama dengan yang diajukan oleh perangkat desa, terutama pada data pekerjaan). Maka saran penulis untuk pengembangan sistem berikutnya adalah dengan melakukan verifikasi NIK dari penyedia data yang algoritmanya dapat mengacu pada Peraturan yang berlaku, sehingga data NIK yang diterima PIC merupakan data NIK yang valid.

Selain itu perlunya peran aktif pemerintah pusat dan daerah yang bekerjasama dengan pemerintah desa / kelurahan sebagai ujung tombak pemerintahan untuk meningkatkan kesadaran masyarakat untuk aktif berperan memperbaharui dan memberikan data identitas yang benar agar tercipta data yang valid dan dapat digunakan dalam segala kegiatan pemerintah untuk mendukung program kesejahteraan masyarakatnya.

## DAFTAR RUJUKAN


Ariestya, W. W., Praptiningsih, Y. E., & Supriatin, W. (2016). Decision Tree Learning Untuk Penentuan Jalur Kelulusan Mahasiswa. *Jurnal Ilmiah FIFO*, *8*(1), 97. https://doi.org/10.22441/fifo.v8i1.1304

*Decision Tree Algorithm Examples In Data Mining*. (n.d.). https://www.softwaretestinghelp.com/decision-tree-algorithm-examples-data-mining/

Durrotul. (2017). Implementasi Data Mining Dalam Prediksi Performance Software Engineer PT.Emerio Menggunakan Decision Tree. *Jurnal Ilmiah Informatika Komputer Universitas Gunadarma*, *22*(1), 31–43.

Kasih, P. (2019). Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara. *Innovation in Research of Informatics (INNOVATICS)*, *1*(2), 63–69. https://doi.org/10.37058/innovatics.v1i2.918

Larose, D. (2005). *Discovering Knowledge in Data : An Introduction to Data Mining*. John Willey & Sons, Inc.

Novianti, T., & Santosa, I. (2016). PENENTUAN JADWAL KERJA BERDASARKAN KLASIFIKASI DATA KARYAWAN MENGGUNAKAN METODE DECISION TREE C4.5 (Studi Kasus Universitas Muhammadiyah Surabaya). *Jurnal Komunika : Jurnal*

*Komunikasi, Media Dan Informatika*, *5*(1), 1. https://doi.org/10.31504/komunika.v5i1.633

Rismayanti. (2018). Decision Tree Penentuan Masa Studi Mahasiswa Prodi Teknik Informatika (Studi Kasus : Fakultas Teknik dan Komputer Universitas Harapan Medan ). *Jurnal Sistem Informasi*, *02*(01), 16–24.

Sutoyo, I. (2018). Implementasi Algoritma Decision Tree Untuk Klasifikasi Data Peserta Didik. *Jurnal Pilar Nusa Mandiri*, *14*(2), 217. https://doi.org/10.33480/pilar.v14i2.926

Turban, E. (2005). *No TitlDecision Support Systems and Intelligent Systems Edisi Bahasa Indonesia Jilid 1e*. Andi Jogjakarta.

Umum, J., Daerah, P., Gresik, K., & Umum, J. (2016). *Bupati gresik provinsi jawa timur*. *3*, 1–20.