
PREVENTIVE DETECTION SYSTEM PADA LAPORAN FRAUD MENGUNAKAN MODEL *TIMESERIES* PADA KANTOR BEA CUKAI

Lela Nurlaela¹⁾, Annisa Qaulan Syadida²⁾

¹⁾Prodi Teknik Informatika, Fakultas Teknologi, ITB Swadharma Jakarta

²⁾Prodi Sistem Informasi, Fakultas Teknologi, ITB Swadharma Jakarta

Correspondence author: L. Nurlaela, lela@swadharma.ac.id, Jakarta, Indonesia

Abstract

State revenue from the excise sector has increased every year because cigarette excise is the biggest contributor to state revenue. Fraud or cheating occurs because of motivation and encouragement from various parties. The techniques of fraud are varied, ranging from outsmarting principles to committing illegal acts were then led to the bankruptcy of the company. According to data from the Directorate General of Customs and Excise, which has its head office in Jalan Jenderal A Yani Rawamangun, East Jakarta, fraud was committed by tobacco senders who did not have excise stamps. The problem in this study is how is the preventive detection system on fraud reports at the customs office. The research objective is to obtain preventive detection measures for fraud reports. The results of time series modeling have a score of 1,006 for the training data and a score of 1 for the data testing showing that the prediction results have a high degree of accuracy.

Keywords : preventive detection, fraud, cigarette excise

Abstrak

Setiap tahun pendapatan negara yang berasal dari sektor cukai mengalami peningkatan karena cukai rokok menjadi penyumbang terbesar bagi pendapatan negara. Fraud atau kecurangan terjadi karena adanya motivasi dan dorongan dari berbagai pihak. Teknik kecurangan atau fraud yang dilakukan bervariasi, mulai dari mengakali prinsip hingga melakukan tindakan ilegal yang kemudian disembunyikan, dan berujung pada kebangkrutan perusahaan. Sesuai data dari Direktorat Jenderal Bea dan Cukai yang berkantor pusat di jalan jenderal A Yani Rawamangun Jakarta Timur, terdapat kecurangan yang dilakukan oleh pihak pengirim tembakau yang tidak terdapat pita cukai. Permasalahan pada penelitian ini yaitu bagaimana *preventive detection system* pada laporan kecurangan di kantor bea cukai? Tujuan Penelitian untuk mendapatkan langkah-langkah *preventive detection* pada laporan kecurangan. Hasil pemodelan *timeseries* dengan nilai skor 1.006 untuk data *training* dan skor 1 untuk data *testing*, menunjukkan bahwa hasil prediksi memiliki tingkat akurasi yang tinggi.

Kata Kunci : *preventive detection*, kecurangan, cukai rokok

A. PENDAHULUAN

Tembakau merupakan salah satu komoditas perdagangan penting di Indonesia. Produk tembakau yang utama diperdagangkan adalah daun tembakau dan rokok. Tembakau dan rokok merupakan produk yang bernilai tinggi, sehingga bagi beberapa negara termasuk Indonesia tembakau dan rokok menjadi salah satu sumber devisa, sumber penerimaan pemerintah dan pajak (Rahman, 2021). Secara ekonomi tembakau dan rokok berdampak positif dan menguntungkan perekonomian khususnya bagi penerimaan negara. Setiap tahun pendapatan negara yang berasal dari sektor cukai mengalami peningkatan karena cukai rokok menjadi penyumbang terbesar bagi pendapatan negara.

Fraud atau kecurangan terjadi karena adanya motivasi dan dorongan dari berbagai pihak, baik dari dalam perusahaan maupun dari luar perusahaan (Komariah, 2016). Dorongan dan motivasi agar laporan yang disajikan terlihat baik dan menarik perhatian. Teknik kecurangan atau *fraud* yang dilakukan bervariasi, mulai dari mengakali prinsip hingga melakukan tindakan ilegal yang kemudian disembunyikan, dan berujung pada kebangkrutan perusahaan. Tidak jarang pula kasus kecurangan laporan yang terjadi, juga melibatkan auditor perusahaan.

Association of Certified Fraud Examiners (ACFE) mendefinisikan *fraud* atau kecurangan yaitu penggunaan suatu jabatan oleh seseorang untuk memperkaya dirinya melalui penyalahgunaan yang disengaja atau penyalahgunaan aset atau sumber daya organisasi. Atau dengan kata lain *fraud* adalah kecurangan berkenaan dengan adanya keuntungan yang diperoleh seseorang dengan menghadirkan sesuatu yang tidak sesuai dengan keadaan yang sebenarnya.

Pesatnya pertumbuhan penerimaan cukai disebabkan oleh kenaikan tarif cukai yang dilakukan pemerintah hampir setiap tahunnya. Berdasarkan peraturan

Kementerian Keuangan nomor 192/PMK.010/2021 menjelaskan bahwa mulai 1 Januari 2022 terdapat kenaikan pada tarif cukai hasil tembakau berupa sigaret, cerutu, rokok daun atau klobot, dan tembakau iris sebesar 12% (Peraturan Menteri Keuangan Republik Indonesia Nomor 125/Pmk.010/2020, 2019). Namun disisi lain, dengan adanya kenaikan terhadap cukai hasil tembakau menyebabkan adanya peningkatan rokok ilegal dan cukai palsu. Jika kecurangan tersebut setiap tahunnya meningkat, maka akan berdampak pada pendapatan negara yaitu penurunan penerimaan dari sektor cukai tembakau.

Sesuai data dari Direktorat Jenderal Bea dan Cukai yang berkantor pusat di jalan jenderal A Yani Rawamangun Jakarta Timur, terdapat kecurangan yang dilakukan oleh pihak pengirim tembakau yang tidak terdapat pita cukai. Oleh karena itu perlu dikaji kembali agar tidak terjadi hal serupa. Sesuai dengan latar belakang tersebut maka diperlukan suatu *preventive detection system* pada laporan kecurangan di kantor bea cukai (Alamsyah, 2011; Suhartono et al., 2015).

Preventive adalah suatu upaya melakukan berbagai tindakan untuk menghindari terjadinya berbagai masalah yang mengancam diri kita sendiri maupun orang lain di masa yang akan datang. *Preventive* adalah usaha pengendalian untuk menghindari atau mencegah sesuatu yang tidak diinginkan. Deteksi adalah proses untuk memeriksa atau melakukan pemeriksaan terhadap sesuatu dengan menggunakan cara dan teknik tertentu.

Sistem adalah sekumpulan elemen-elemen yang berinteraksi dengan maksud yang sama untuk mencapai suatu tujuan tertentu (Sutabri, 2012). Untuk memperluas wawasan dan pendalaman materi dilakukan juga studi dari beberapa penelitian terdahulu antara lain. Penelitian yang dilakukan oleh Adhitio Satyo Bayangkari (Bayangkari Karno, 2020). Penelitian ini menggunakan metode statistik ARIMA yang ditentukan melalui nilai parameter dari *Auto-Regressive*

(p), *Intergrated* (d), *Moving Average* (q) melalui grafik *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) dan metode mesin belajar yang mampu mengolah data runtun waktu yaitu *Long Short Term Memory* (LSTM), dengan melakukan beberapa proses awal (*cleansing, transformasi, smooting dan differencing*). Dengan harapan dari kombinasi proses awal pada kedua metode ini akan menghasilkan tingkat akurasi yang lebih baik. Dan penggunaan bahasa pemrograman *Python* karena bahasa ini mempunyai banyak keunggulan khususnya untuk pemrograman berbasis *machine learning*. Hasil penelitian ini menyebutkan bahwa prediksi data *time series* memperlihatkan mesin belajar LSTM lebih akurat dibandingkan dengan menggunakan model statistik ARIMA. Pemilihan proses transformasi yang sesuai untuk model ARIMA dan LSTM masih perlu dikembangkan agar memperoleh hasil yang lebih baik lagi.

Menurut Fauzah Umami, Hendra Cipta, Ismail Husein (Umami et al., 2019). Penelitian ini adalah jenis penelitian dengan pendekatan kuantitatif yaitu dengan mengambil dan mengumpulkan data mentah sesuai kebutuhan, dan menganalisisnya dengan data *time series*. Dalam analisis penelitian ini digunakan *time series* dengan metode peramalan yaitu *forecasting*. Penelitian ini data diolah dengan menggunakan software IBM SPSS 22 dan MINITAB 19. Analisis data yang digunakan yaitu mengumpulkan data, mendeskripsikan data, menganalisis data, uji stasioneritas, penggunaan model ARIMA, menafsirkan hasil analisis dan menjelaskan model yang terbentuk dari hasil peramalan. Hasil penelitian model analisis data *time series* yang terbaik dengan menggunakan model ARIMA (0,2,3) dengan *Mean Square Error* (MSE) model adalah 155,182 dan derajat kebebasan (df) dari model adalah 25 dan persamaan model sebagai berikut:

$$Z_t = (1 + \theta_1)Z_{t-1} + \theta Z_{t-2} + a_t + \theta_1 a_{t-1}$$

$$Z_t = (1+1,037)Z_{t-1} + 1,037Z_{t-2} + a_t + 0,1423a_{t-1}$$

Menurut Anis Mahfud Al Afi, Widiarti, Dian Kurniasari, dan Mustofa Usman (Al'afi et al., 2020). Penelitian ini menggunakan pemodelan SARIMA dengan mencari periodesitas tersembunyi dari data menggunakan analisis spektral. Sebelum melakukan analisis spektral, data diuji kestasionerannya melalui uji ADF. Jika data tidak stasioner maka data perlu ditransformasi dan *differencing* agar menjadi stasioner. Kemudian data diuji Kembali dengan uji ADF. Hasil penelitian yaitu model SARIMA (0,1,1)(0,1,1)³ merupakan model terbaik dari deret waktu untuk meramalkan dengan persamaan model sebagai berikut:

$$Y_t = 1,6457Y_{t-1} - 0,6457Y_{t-2} + 0,3543Y_{t-3} - Y_{t-4} + 0,6457Y_{t-5} + e_t - 0,9114e_{t-3}$$

Tujuan Penelitian untuk mendapatkan langkah-langkah preventive detection pada laporan kecurangan.

B. METODE PENELITIAN

Metode penelitian yang digunakan pada penelitian ini menggunakan metode studi lapangan dengan teknik pengumpulan data yaitu observasi dan metode studi pustaka.

Observasi pengumpulan data terhadap laporan fraud pada kantor bea cukai cabang Rawamangun pada tanggal 1 Januari 2021 sampai 31 Juli 2021.

Studi Pustaka untuk melakukan kajian teoritis, referensi serta literatur ilmiah lainnya yang berkaitan dengan budaya, nilai dan norma yang berkembang pada situasi sosial yang diteliti.

Mekanisme pengumpulan data yang dilakukan bea cukai pada saat ini sudah cukup bagus namun penindakan terhadap kasus cukai ilegal masih perlu ditingkatkan. Berdasarkan uraian pada bab sebelumnya, *Preventive detection* dikembangkan dari data penindakan. Maka dalam pengembangan dilakukan langkah sebagai berikut:



Gambar 1. Mekanisme Analisis.

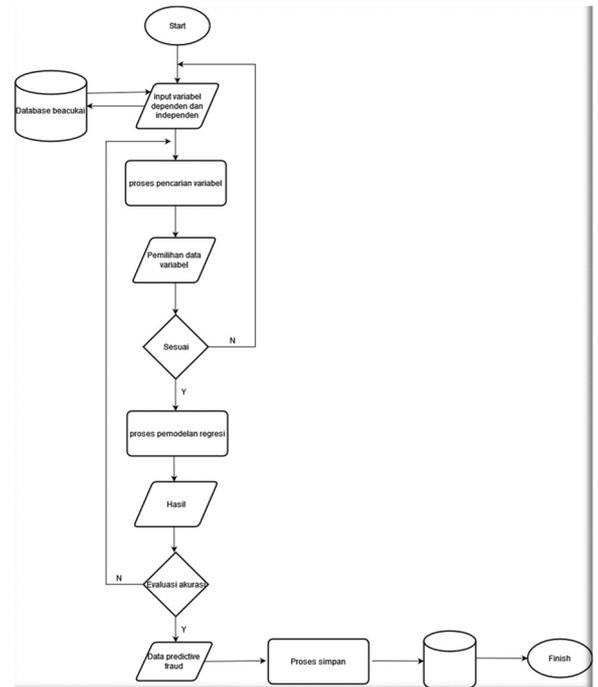
C. HASIL DAN PEMBAHASAN

Pada tahap awal pengembangan *preventive detection* dilakukan input data ke dalam *interface development Environment*. Semua *field* dan semua data dimasukkan dalam *node input* data. Kemudian dilakukan proses data *preparation* dan pemodelan data sehingga hasil akhirnya didapat data model.

Pengembangan Preventive Detection

Pengembangan *preventive detection* dilakukan dalam beberapa tahap sebagai berikut: *Business Understanding* seperti terlihat pada Gambar 2. Pada tahap ini ditentukan *GOAL* atau tujuan pengembangan. Adapun yang menjadi tujuan/*GOAL* adalah menemukan angka prediksi jumlah kenaikan ataupun penurunan pada periode selanjutnya yang masih akan dijalani. Sehingga dengan angka prediksi tersebut, maka bisa dilakukan tindakan *preventive* jika terjadi kenaikan jumlah kasus. Tindakan *preventive* yang dimaksud adalah tindakan yang perlu dilakukan sebagai berikut: Mengantisipasi kenaikan jumlah cukai ilegal. Antisipasi dilakukan seperti pembagian tim kerja berdasarkan prediksi kenaikan jumlah kejadian. Sosialisasi kasus. Hal ini perlu dilakukan karena lonjakan kasus bisa saja terjadi karena ketidaktahuan masyarakat terhadap kasus-kasus pelanggaran. Sosialisasi dilakukan pada daerah tertentu dan tanggal-tanggal tertentu yang di prediksi naik.

Pemahaman Data (*Data Understanding*). Tahap ini dilakukan untuk memahami data. Data yang dijadikan objek pengembangan adalah data penindakan dikumpulkan dalam kurun waktu 1 Januari 2021 sampai 31 Juli 2021 yang terdiri dari 58 *field* dan 19653 baris.



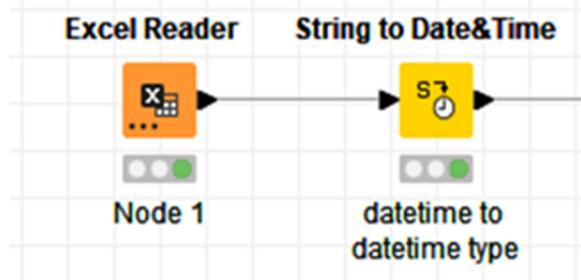
Gambar 2. Mekanisme Preventive Detection

Persiapan Data (*Data preparation*). Tahap ini dilakukan proses *cleansing* dan *handling missing value*. Dari *observasi* data ditemukan *field longitude* dan *latitude* masih kosong. Dalam upaya untuk mengisinya, maka dilakukan pengisian data *longitude* dan *latitude* berdasarkan kabupaten dan kota.

Pemodelan data (*Data Modelling*). Identifikasi permasalahan *preventive detection* adalah permasalahan prediksi yang termasuk dalam kelompok *supervised learning*. pengolahan data dengan memodelkan data untuk mendapatkan pola data *timeseries*. Pola data *timeseries* digunakan untuk menemukan pola kenaikan atau penurunan jumlah kasus yang disebut *fraud*. Adapun untuk proses model, data akan dibagi menjadi 2 yaitu:

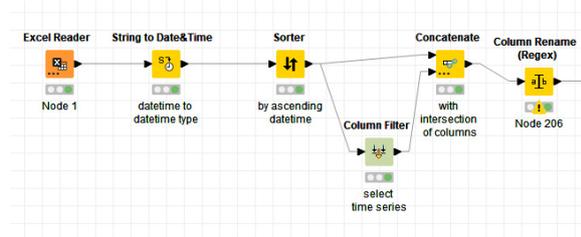
Data training sebanyak 70% dari total data dimulai dari tanggal 1 Januari 2021 sampai dengan tanggal 31 Mei 2021 diambil dari kolom tanggal penindakan. Data testing sebanyak 30% dari total data dimulai dari tanggal 1 Juni 2021 sampai dengan tanggal 31 Juli 2021 diambil dari kolom tanggal penindakan. Untuk variabel yang dipilih

pada proses pemodelan adalah sebagai berikut: Tanggal penindakan, Jumlah kasus perhari, Perkiraan nilai barang. Selanjutnya pengolahan atau pemodelan data training menggunakan knime *analytic tools* sebagai berikut:



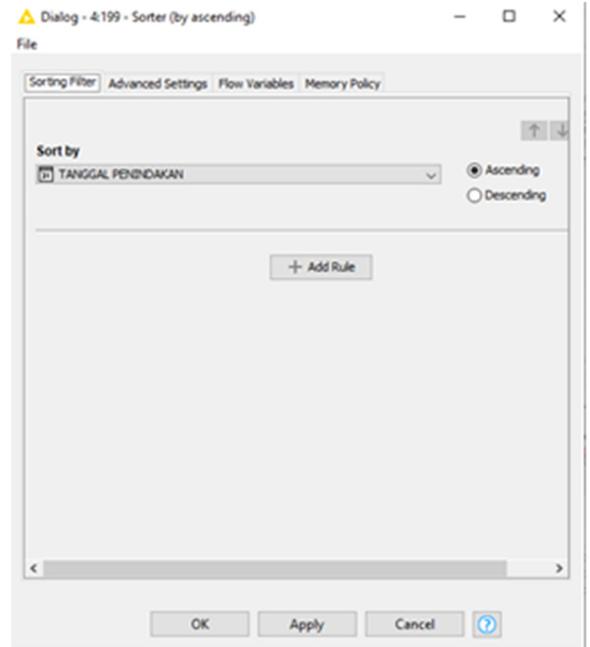
Gambar 3. Proses Input dan Pemilihan Variabel

Pada tahap awal dilakukan input data. Data diperoleh dari sumber data langsung yang berupa file excel. Pada *software* knime untuk proses awal digunakan node excel reader untuk membaca file excel yang diperoleh, dan node *string to date&time* digunakan untuk merubah *field* yang bertipe data *string* diubah menjadi *date&time*.

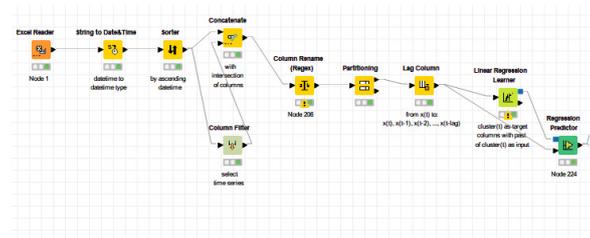


Gambar 4. Persiapan Data

Selanjutnya penggunaan node *sorter* untuk mengurutkan kolom yang datanya harus diurutkan. Pada gambar 5 tersebut kolom yang dipilih tanggal penindakan pemilihan kolom tersebut karena kolom tersebut bertipe data *date*. Node selanjutnya yaitu *column filter* yang berfungsi untuk memilih kolom mana yang akan diproses pada node selanjutnya.

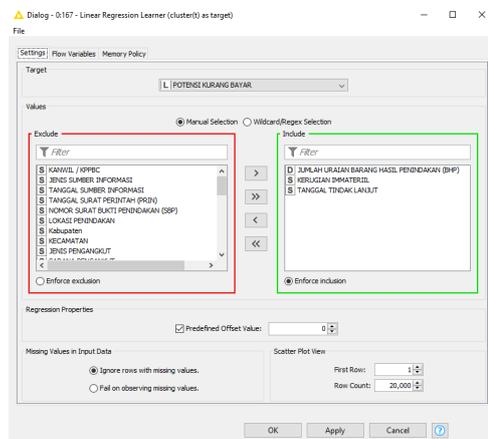


Gambar 5. Proses Sorter



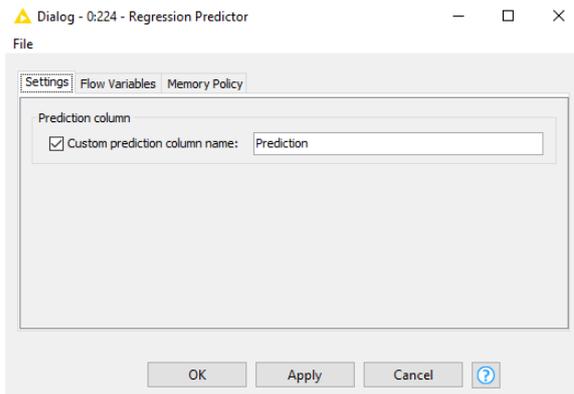
Gambar 6. Proses Pemodelan

Proses pemodelan dilakukan menggunakan node *linear regression learner* kemudian dilanjutkan dengan node *regression predictor*.



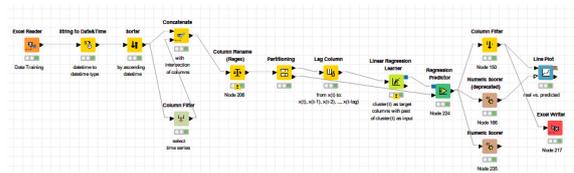
Gambar 7. Linear Regression Learner

Pada node *linear regression learner* dilakukan pemilihan target variabel. Variabel target yang dipilih yaitu potensi kurang bayar, dan variabel *include* yang dipilih yaitu jumlah uraian barang hasil penindakan (BHP), kerugian immateriil, dan tanggal tindak lanjut.

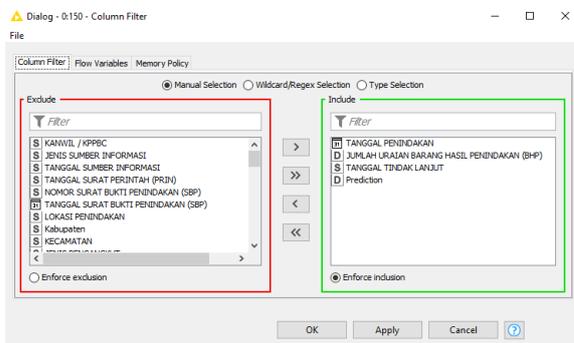


Gambar 8. Regression Predictor

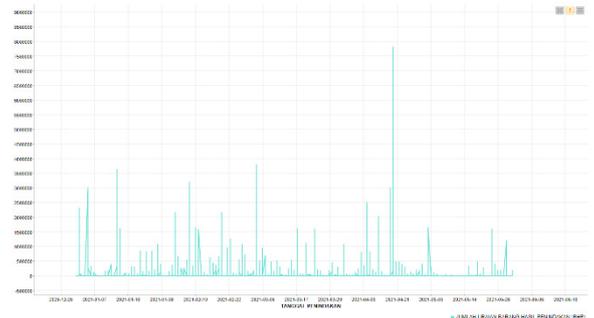
Kemudian pada node *regression predictor* pembuatan kolom baru yaitu *prediction*.



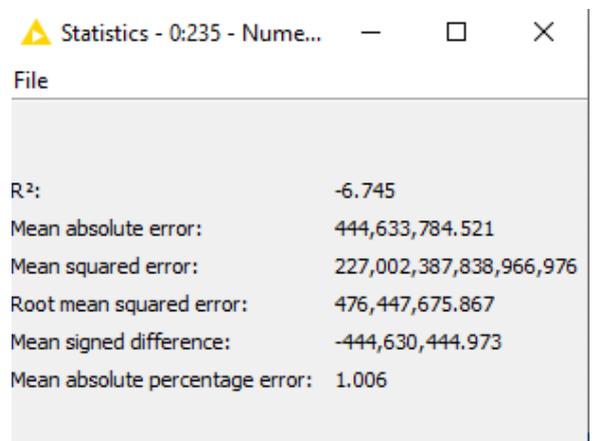
Gambar 9. Workflow Pada Data Training



Gambar 10. Pemilihan Variabel Pada Column Filter



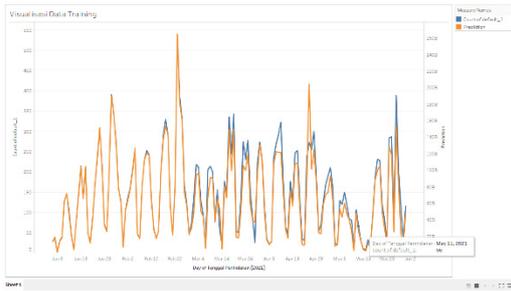
Gambar 11. Output Line Plot



Gambar 12. Output Numeric Scorer

Dari hasil node numeric scorer dapat dilihat pada *mean absolute percentage error* yaitu 1.006 yang menyatakan bahwa hasil rata-rata akurasi prediksi kesalahan pada data training akurat.

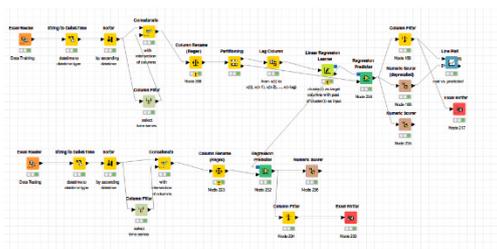
Gambar 13. Output Excel Writer Data Training



Gambar 14. Hasil Pemodelan Prediction VS Actual

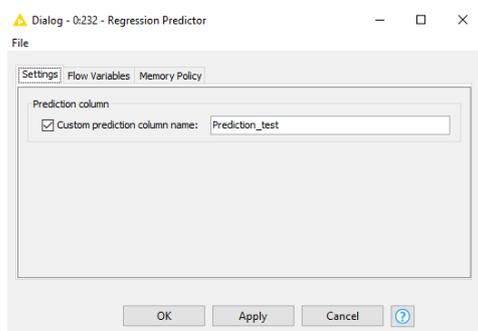
Pada gambar diatas hasil pemodelan terhadap *prediction vs actual* menggambarkan bahwa hasil prediksi tidak jauh berbeda dengan aktual karena persentasi hasil rata-rata akurasi 1.006.

Kemudian dilakukan pemodelan kembali terhadap data testing, sebagai berikut:



Gambar 15. Workflow Data Testing

Proses yang dilakukan pada data testing hampir sama seperti data training. Dimulai dengan node *excel reader* hingga *excel writer*. Node *excel reader* pada data testing tentunya menggunakan data testing.



Gambar 16. Regression Predictor Data Testing

Kemudian pada node *regression predictor* pembuatan variabel baru yaitu *prediction_test* yang akan digunakan pada grafik visualisasi.

Statistics - 0:236 - Numeric Scorer

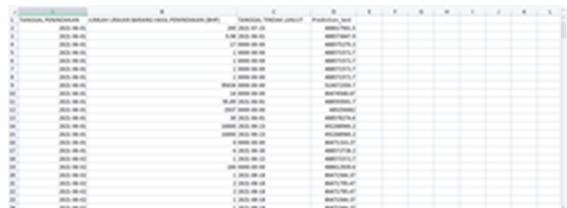
File Edit Hilite Navigation View

Table "Scores" - Rows: 6 Spec - Column: 1 Properties Flow Variables

Row ID	D JUMLA...
R^2	-5.516
mean absolute error	401,285,01...
mean squared error	190,218,94...
root mean squared error	436,140,96...
mean signed difference	-401,285,01...
mean absolute percentage error	1

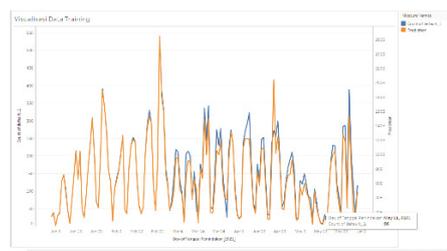
Gambar 17. Output Node Number Scorer

Dapat dilihat pada gambar diatas, nilai *mean absolut percentage error* pada data testing yaitu 1 membuktikan bahwa data training dengan data testing memiliki hasil yang akurat.



Gambar 18. Output Excel Data Testing

Dari analisa dapat disimpulkan bahwa *preventive detection* belum dilakukan pada kasus *fraud* oleh karena itu dilakukan pemodelan terhadap data dengan pemodelan *timeseries* dengan nilai *scorer* 1.006 untuk data *training* dan *scorer* 1 untuk data *testing*. Hasil tersebut menunjukkan bahwa hasil prediksi memiliki tingkat akurasi yang tinggi. Berikut hasil visualisasi data *training* dan data *testing*:



Gambar 19. Hasil Pemodelan Prediction vs Actual



Gambar 20. Hasil Prediction 6 Bulan Yang Akan Datang

Pada gambar diatas merupakan gambaran hasil prediction kasus cukai ilegal yang akan terjadi untuk 6 bulan yang akan datang. Maka rekomendasi preventive detection untuk hasil prediction tersebut yaitu dilakukan mapping man power pada maksimum capaian prediksi kasus untuk mengatasi lamanya waktu penindakan. Dengan dilakukannya mapping man power diharapkan penindakan dapat dilakukan secara efektif. Kemudian diperoleh saran rekomendasi preventive detection untuk kasus fraud.

Rekomendasi Hasil Preventive Detection

Berikut rekomendasi preventive detection dari hasil prediction:

1. Mengantisipasi kenaikan jumlah cukai ilegal.
2. Antisipasi dilakukan seperti pembagian tim kerja berdasarkan prediksi kenaikan jumlah kejadian.
3. Sosialisasi kasus. Hal ini perlu dilakukan karena lonjakan kasus bisa saja terjadi karena ketidaktahuan masyarakat terhadap kasus-kasus pelanggaran.

Sosialisasi dilakukan pada daerah tertentu dan tanggal-tanggal tertentu yang di prediksi naik.

D. PENUTUP

Dari analisa tersebut dapat disimpulkan bahwa preventive detection penting dilakukan untuk mengantisipasi kasus fraud.

Pembuatan alarm warning untuk kasus fraud sehingga berdampak pada penindakan

mapping man power dan dilakukan evaluasi model preventive secara berkala. Pengembangan lebih lanjut dengan mengevaluasi hasil analisa dan disesuaikan dengan kebutuhan.

E. DAFTAR PUSTAKA

- Al'afi, A. M., Widiart, W., Kurniasari, D., & Usman, M. (2020). Peramalan Data Time Series Seasonal Menggunakan Metode Analisis Spektral. *Jurnal Siger Matematika*, 1(1), 10–15. <https://doi.org/10.23960/jsm.v1i1.2484>
- Alamsyah. (2011). Implementasi Keamanan Intrusion Detection System (IDS) dan Intrusion Prevention System (IPS) Menggunakan Clearos. *Jurnal SMARTek*, 9(3), 223–229.
- Bayangkari Karno, A. S. (2020). Analisis Data Time Series Menggunakan LSTM (Long Short Term Memory) Dan ARIMA (Autocorrelation Integrated Moving Average) Dalam Bahasa Python. *Ultima InfoSys: Jurnal Ilmu Sistem Informasi*, 11(1), 1–7. <https://doi.org/10.31937/si.v9i1.1223>
- Karaa, F. A. (1989). Infrastructure maintenance management system development. *Journal of Professional Issues in Engineering Education and Practice*, 115(4), 422–432. [https://doi.org/10.1061/\(ASCE\)1052-3928\(1989\)115:4\(422\)](https://doi.org/10.1061/(ASCE)1052-3928(1989)115:4(422))
- Komariah, D. (2016). *Security Management Control Pada Jaringan Komputer*. Fakultas Ilmu Komputer Universitas Sriwijaya.
- Peraturan Menteri Keuangan Republik Indonesia Nomor 125/Pmk.010/2020, 15 1 (2019).
- Rahman, H. A. (2021). *Agresivitas Pajak dan Faktor-faktor yang Memengaruhinya*. 6(2), 195–206.
- Sugiyono. (2017). *Metode Penelitian Bisnis*:

Pendekatan Kuantitatif, Kualitatif, Kombinasi dan R&D. Bandung : Alfabeta.

Suhartono, D., Riyanto, A. D., & Astomo, Y. W. (2015). Intrusion Detectin Prevention System (IDPS) Pada Local Area Network (LAN). *Telematika*, 8(1), 24–42.

Sutabri, T. (2012). *Analisa Sistem Informasi*. Andi Offset.

Umami, F., Cipta, H., & Husein, I. (2019). Data Analysis Time Series For Forecasting The Greenhouse Effect. *ZERO: Jurnal Sains, Matematika Dan Terapan*, 3(2), 86. <https://doi.org/10.30829/zero.v3i2.7914>