

Prediksi Jumlah Tunggakan Pajak Jatuh Tempo Menggunakan Algoritma *Support Vector Regression*

Mustakim¹, Celsa Bella², Yoga Rizola Pratama³

Program Studi Sistem Informasi Fakultas Sains dan Teknologi^{1,2,3}
Puzzle Research Data Technology^{1,2,3}

Universitas Islam Negeri Sultan Syarif Kasim Riau

Jl. HR. Soebrantas KM 18 No. 155 Pekanbaru Riau Indonesia

e-mail: mustakim@uinsuska.ac.id¹; celsabella1997@gmail.com²; yogarizola06@gmail.com³.

Abstrak

Pajak kendaraan bermotor merupakan pajak daerah yang memiliki kontribusi besar bagi peningkatan pendapatan asli daerah. Salah satu sumber pendapatan asli daerah di Kabupaten Pasaman adalah dana bagi hasil pajak kendaraan, yang dialokasikan untuk pendidikan, pembangunan dan sarana transportasi. Peningkatan tunggakan pajak kendaraan berpengaruh signifikan terhadap pendapatan asli daerah dan akan mengganggu kelangsungan pembangunan daerah. Penelitian ini dimaksudkan untuk melakukan prediksi jumlah tunggakan pajak kendaraan jatuh tempo di Kabupaten Pasaman dengan penerapan data mining yaitu algoritma *support vector regression* sehingga akan diketahui jumlah tunggakan untuk periode selanjutnya. Hasil prediksi periode pertama menunjukkan sebanyak 8 Kecamatan mengalami kenaikan jumlah tunggakan pajak, dan pada periode kedua sebanyak 7 Kecamatan mengalami peningkatan jumlah tunggakan. Model prediksi terbaik menggunakan teknik *k-Fold Cross Validation*, untuk periode pertama didapat pada *Fold-1* dengan nilai MSE 4,8% dan Tingkat Akurasi (R) 91%, untuk periode kedua model terbaik juga didapat pada *Fold-1* dengan MSE 0,14% dan R 96%.

Kata kunci: Data Mining, Pajak Kendaraan Bermotor, Prediksi, *Support Vector Regression*.

Abstract

Motor vehicle tax is a regional tax which has a major contribution to increase in local revenue. One source of local revenue in Pasaman Regency is the vehicle tax revenue sharing fund, which is allocated for education, development and transportation facilities. The increase in vehicle tax arrears has a significant effect on local revenue and will disrupt the continuity of regional development. This study is intended to predict the amount of vehicle tax arrears due in Pasaman by applying data mining, namely the *support vector regression* algorithm so that the arrears will be known for the next period. The prediction result of the first period showed that 8 Sub-districts experienced an increase in the amount of tax arrears, and in the second period there were 7 Sub-Districts which experienced an increase in the arrears amount. The best prediction model uses the *k-Fold Cross Validation* technique, for the first period obtained in *Fold-1* with a value of 4.8% MSE and Accuracy Rate (R) 91%, for the second period the best model is also obtained in *Fold-1* with MSE 0, 14% and R 96%.

Keywords: Data Mining, Motor Vehicle Tax, Prediction, *Support Vector Regression*.

1. Pendahuluan

Pajak adalah kontribusi wajib yang harus dipenuhi kepada negara yang terutang oleh pribadi atau badan dan bersifat memaksa berdasarkan undang-undang dengan tidak mendapatkan imbalan secara langsung yang digunakan untuk keperluan negara bagi sebesar-besarnya kemakmuran rakyat (UU No. 28 tahun 2007 Tentang KUP). Pajak Kendaraan Bermotor merupakan salah satu jenis pajak daerah yang memberikan penerimaan dan kontribusi cukup besar dalam meningkatkan Pendapatan Asli Daerah [1]. Unit Pelayanan Teknis Daerah Pengelolaan Pendapatan Daerah (UPTD-PDD) Samsat Kabupaten Pasaman merupakan salah satu lembaga daerah yang melaksanakan pelayanan pajak, meliputi pajak kendaraan bermotor, balik nama kendaraan bermotor, serta penagihan pajak. Pendapatan asli daerah yang memberikan kontribusi terbesar bagi pemerintah daerah Kabupaten Pasaman adalah penerimaan sektor lain-lain yang di dalamnya termasuk dana bagi hasil pajak kendaraan bermotor sebesar 35.594 milyar rupiah. Pemanfaatan dana bagi hasil pajak kendaraan dialokasikan untuk pembangunan, pemeliharaan jalan, sarana transportasi dan pendidikan [2].

Data yang dihimpun dari PPD Lubuk Sikaping, pada tahun 2017 jumlah Kendaraan Bermotor mencapai 24.251 unit, jumlah ini sedikit mengalami penurunan dari tahun sebelumnya yaitu 25.128 unit pada tahun 2016. Sementara itu, untuk data analisa kendaraan tidak daftar ulang (jatuh tempo) tahun 2017 tercatat lebih dari 5.728 unit kendaraan bermotor tidak melakukan daftar ulang, dengan total pajak mencapai 1,4 miliar rupiah, jumlah ini mengalami penurunan dari tahun

sebelumnya, dengan total pajak 2 miliar rupiah dan jumlah kendaraan 4.178 unit pada tahun 2016. Banyaknya tunggakan pajak kendaraan yang disebabkan tidak mendaftarkan ulang dan jatuh tempo berakibat terhadap pendapatan daerah dan akan mengganggu kelangsungan pembangunan di daerah tersebut. Salah-satu kendala yang dapat menghambat keefektifan pengumpulan pajak adalah kepatuhan wajib pajak (*tax compliance*) [3].

Melihat besarnya kontribusi pajak kendaraan bermotor sebagai salah satu pendapatan asli daerah, tentunya pendapatan asli daerah diharapkan untuk meningkat setiap tahunnya, dengan meminimalkan pajak kendaraan yang jatuh tempo serta meningkatkan kesadaran dan kepatuhan wajib pajak dalam melaksanakan kewajibannya. Oleh sebab itu, pada penelitian ini akan dilakukan prediksi jumlah tunggakan pajak kendaraan jatuh tempo dengan teknik data mining, sehingga akan diketahui jumlah tunggakan pajak di masa mendatang untuk dapat diproyeksikan kepada Pendapatan asli daerah di Kabupaten Pasaman dan juga sebagai evaluasi bagi pemerintah dalam proses pengambilan keputusan untuk meningkatkan kesadaran wajib pajak.

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstraksi pengetahuan secara otomatis [4]. *Data mining* berisi pencarian pola atau *trend* yang dibutuhkan dalam pengambilan sebuah keputusan pada masa yang akan datang [5]. Salah satu algoritma yang dapat menyelesaikan kasus prediksi adalah *Support Vector Regression* (SVR) dengan akurasi yang lebih baik dibandingkan metode lain seperti *Artificial Neural Network* [6]. Metode SVR telah diterapkan secara luas dari permasalahan kehidupan nyata dan berfungsi secara sempurna dalam beberapa kasus dengan menghasilkan kinerja regresi secara akurat [7].

Teknik prediksi dibutuhkan untuk mengetahui gambaran keadaan masa depan, penelitian dengan teknik prediksi menggunakan algoritma SVR dilakukan oleh Agustina (2018), yaitu pemodelan algoritma SVR untuk prediksi ketersediaan bahan pangan di Indonesia dalam menghadapi bonus Demografi diperoleh model terbaik untuk hasil prediksi menggunakan kernel *Radial Basis Function* pada SVR dengan koefisien determinasi (R) 92% dengan nilai error galat 14% [8].

2. Metodologi Penelitian

Data yang digunakan untuk prediksi dalam Penelitian ini ialah data tunggakan pajak kendaraan bermotor se-Kecamatan yang ada Kabupaten Pasaman dari tahun 2015-2018, di mana terdapat 12 Kecamatan dengan pelabelan nama dikarenakan data bersifat rahasia. Data yang akan diprediksikan adalah jumlah tunggakan setiap Kecamatan untuk dua periode yaitu periode pertama Januari-Maret, dan periode kedua April-Juni 2019.

Tahapan pengolahan data dengan penerapan Algoritma SVR di antaranya, normalisasi data, pembagian data training dan testing dengan teknik *K-fold cross validation*, optimasi dengan *grid search* dan kernel *Radial Basis Function* yang dapat dilihat pada Gambar 1.

2.1. Pajak Kendaraan Bermotor

Pajak Kendaraan Bermotor merupakan pajak terhadap kepemilikan atau penguasaan kendaraan bermotor baik kendaraan roda dua atau lebih (UU No. 28 Tahun 2009 tentang Pajak dan Retribusi Daerah). Dana Hasil pajak kendaraan difokuskan untuk pendidikan, infrastruktur jalan, memelihara dan meningkatkan kesehatan serta mencegah dan memberantas penyakit.

2.2. Data Mining

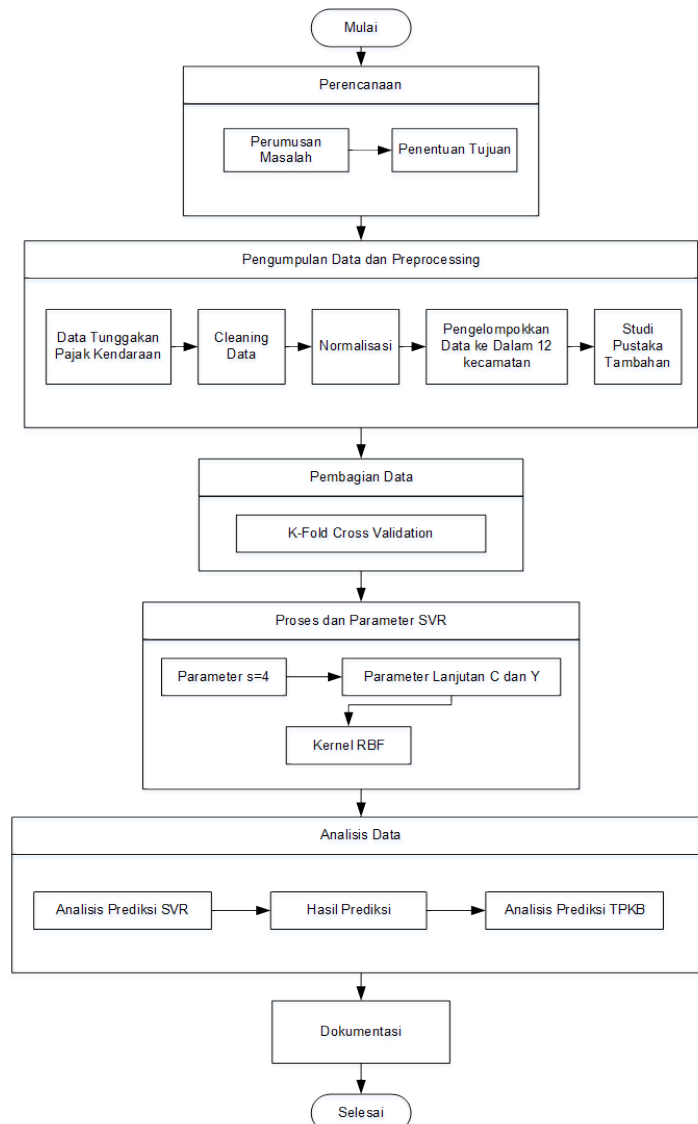
Data Mining merupakan proses menganalisis data mentah dengan bantuan komputer untuk mengekstraksi makna dari kumpulan data. Hal ini sering didefinisikan sebagai menemukan informasi dari data yang sebelumnya tidak diketahui sehingga memperoleh potensi dan manfaat dari volume data besar [9]. Teknik data mining dapat membantu pemilik data untuk menganalisis dan menemukan hubungan yang tidak terduga di antara data yang nantinya dapat digunakan sebagai pendukung keputusan [10].

2.3. Support Vector Regression

Support vector regression adalah aplikasi dari *Support Vector Machine* (SVM) untuk kasus-kasus regresi, pada kasus regresi, *output* yang dihasilkan dalam bentuk bilangan nyata (*riil*) [11]. SVR adalah metode yang dapat mengatasi *overfitting*, sehingga akan menghasilkan kinerja yang baik [12], [13]. Algoritma SVR mempunyai keunggulan dibandingkan dengan model regresi linier berganda (RLB). Keunggulannya terletak dalam hal pemanfaatan data non linier secara implisit melalui penerapan fungsi kernel [8].

2.4. K- Fold Cross Validation

K-fold cross validation adalah metode praktis yang populer untuk mendapatkan perkiraan yang baik dari tingkat kesalahan dari algoritma pembelajaran [14]. *K-fold cross validation* digunakan untuk membagi data menjadi *training set* dan *test set*, inti dari validasi tipe ini adalah membagi data secara acak ke dalam k himpunan bagian yang diinginkan [12].



Gambar 1. Metodologi Penelitian

3. Hasil dan Analisis

Data yang digunakan pada penelitian ini bersumber dari Samsat Kabupaten Pasaman, yaitu data jumlah tunggakan pajak kendaraan jatuh tempo di seluruh Kecamatan yang ada di Kabupaten Pasaman dari tahun 2015-2018. Data yang akan diprediksikan adalah jumlah tunggakan setiap Kecamatan untuk dua periode yaitu Januari-Maret, April-Juni 2019. Kecamatan yang ada digunakan sebagai Atribut dan jumlah data yang digunakan sebanyak 12 data sesuai dengan jumlah Kecamatan yang ada di Pasaman.

3.1. Normalisasi Data

Normalisasi data bertujuan untuk menghasilkan nilai keseimbangan antara nilai yang rendah dengan nilai yang tinggi, dengan cara mencari data awal, data maksimum dan minimum. Hasil normalisasi data dapat dilihat pada Tabel 1.

Tabel 1. Hasil Normalisasi Data

| Kecamatan | 1 | 2 | ... | 15 | 16 | Aktual |
|-----------|-------|-------|-----|-------|-------|--------|
| A | 0,256 | 0,404 | | 0,213 | 0,196 | 0,268 |
| B | 0,734 | 0,957 | | 1,000 | 1,000 | 0,944 |
| C | 0,514 | 0,737 | | 0,425 | 0,378 | 0,510 |
| D | 0,161 | 0,171 | | 0,009 | 0,013 | 0,102 |
| E | 0,075 | 0,110 | | 0,096 | 0,222 | 0,116 |
| F | 0,522 | 0,841 | | 0,517 | 0,461 | 0,745 |
| G | 1,000 | 1,000 | | 0,393 | 0,409 | 0,685 |
| H | 0,011 | 0,066 | | 0,000 | 0,000 | 0,011 |
| I | 0,072 | 0,024 | | 0,027 | 0,083 | 0,068 |
| J | 0,452 | 0,349 | | 0,217 | 0,284 | 0,345 |
| K | 0,000 | 0,000 | | 0,042 | 0,064 | 0,018 |
| L | 0,096 | 0,173 | | 0,189 | 0,190 | 0,193 |

3.2. Pembagian data dengan k-fold cross validation

K-fold cross validation digunakan untuk membagi data menjadi *training* set dan *test* set, inti dari validasi tipe ini adalah membagi data secara acak ke dalam k himpunan bagian yang diinginkan. Pada penelitian ini data *training* dan data *testing* dibagi dengan teknik *K-Fold Cross Validation* sebanyak 4 *Fold*.

3.3. Optimasi dengan Grid Search

Grid Search mengoptimalkan parameter *SV-Machine* (*C*, *gamma*, *degree*) menggunakan teknik *cross validation* sebagai metrik kinerja. Penggunaan optimasi dengan *Grid Search* mampu menghasilkan nilai yang optimal. Optimasi *grid search* ini digunakan secara *default* pada Matlab dengan LIBSVM. Pada penelitian ini kernel yang digunakan ialah *Radial Basis Function* (RBF) dengan parameter *C* dan *Gamma*, uji parameter *C* dan *Gamma* akan menghasilkan nilai MSE dan R dan hasil prediksi. Tabel 2 merupakan hasil penguji kernel RBF untuk *fold*-1. Untuk *fold* 2, 3 dan 4 pengujian dilakukan dengan cara yang sama.

Tabel 2. Hasil Percobaan Kernel RBF pada *fold*-1

| γ | C | | 0,25 | 0,5 | 0,75 | 1 |
|----------|---|-----|--------|--------|--------|--------|
| | | | 0,5 | 0,75 | 1 | 2 |
| 0.5 | 1 | MSE | 0,0758 | 0,0802 | 0,0870 | 0,0911 |
| | | R | 0,9163 | 0,9163 | 0,8891 | 0,8730 |
| 0.25 | 2 | MSE | 0,0783 | 0,0822 | 0,0874 | 0,0951 |
| | | R | 0,9103 | 0,9092 | 0,9054 | 0,8577 |
| 0.125 | 4 | MSE | 0,0408 | 0,0408 | 0,0408 | 0,0408 |
| | | R | 0,6153 | 0,6153 | 0,6153 | 0,6153 |
| 0.0625 | 8 | MSE | 0,0552 | 0,0552 | 0,0552 | 0,0552 |
| | | R | 0,3587 | 0,3587 | 0,3587 | 0,3587 |

Nilai MSE merupakan ukuran error pada percobaan, nilai terbaik adalah MSE terkecil mendekati 0 dan nilai Akurasi terbesar mendekati 1. Model terbaik dari nilai MSE dan R yang nantinya direkomendasikan menjadi Prediksi pilihan.

3.4. Hasil Prediksi Terbaik untuk Periode Pertama (Januari-Maret 2019)

Untuk hasil prediksi jumlah tunggakan pajak masing-masing kecamatan pada periode pertama bulan Januari-Maret 2019 dipilih *fold* yang mempunyai tingkat akurasi terbaik $R=0,9163$ dan $MSE = 0,0408$ yang diperoleh dari *fold*-1. Berikut hasil prediksi periode pertama yang dapat dilihat pada Tabel 3.

Hasil prediksi periode pertama akan dibandingkan dengan nilai Aktual yang merupakan nilai rata-rata setiap periode pertama 4 tahun terakhir. Perbandingan ini bertujuan untuk melihat *trend* kenaikan atau penurunan jumlah tunggakan pajak kendaraan bermotor. Perbandingan dapat dilihat pada Tabel 4.

Tabel 3. Hasil Prediksi Periode Pertama setelah di Denormalisasi.

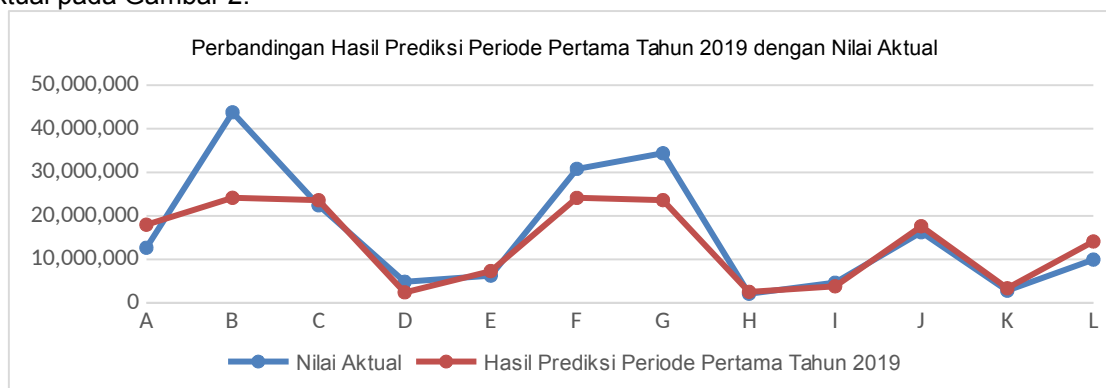
| N | Nama Kecamatan | Prediksi Periode Pertama | Hasil Prediksi setelah Denormalisasi (Rp) |
|---|----------------|--------------------------|---|
|---|----------------|--------------------------|---|

| No | Nama Kecamatan | Nilai Aktual | Hasil Prediksi Periode Pertama Tahun 2019 | Trend Hasil Prediksi |
|----|----------------|--------------|---|----------------------|
| 1 | A | 0,244 | 17,965,851 | Naik |
| 2 | B | 0,340 | 24,143,952 | Turun |
| 3 | C | 0,331 | 23,594,367 | Naik |
| 4 | D | 0,002 | 2,426,534 | Naik |
| 5 | E | 0,078 | 7,317,173 | Naik |
| 6 | F | 0,340 | 24,124,126 | Turun |
| 7 | G | 0,332 | 23,605,989 | Turun |
| 8 | H | 0,004 | 2,533,502 | Naik |
| 9 | I | 0,023 | 3,811,740 | Turun |
| 10 | J | 0,238 | 17,611,842 | Naik |
| 11 | K | 0,016 | 3,357,161 | Naik |
| 12 | L | 0,184 | 14,110,933 | Naik |

Tabel 4. Perbandingan Hasil Prediksi Periode Pertama dengan Nilai Aktual

| No | Nama Kecamatan | Nilai Aktual | Hasil Prediksi Periode Pertama Tahun 2019 | Trend Hasil Prediksi |
|----|----------------|--------------|---|----------------------|
| 1 | A | 12,676,877 | 17,965,851 | Naik |
| 2 | B | 43,782,100 | 24,143,952 | Turun |
| 3 | C | 22,412,690 | 23,594,367 | Naik |
| 4 | D | 4,855,286 | 2,426,534 | Naik |
| 5 | E | 6,262,830 | 7,317,173 | Naik |
| 6 | F | 30,790,033 | 24,124,126 | Turun |
| 7 | G | 34,381,081 | 23,605,989 | Turun |
| 8 | H | 2,103,803 | 2,533,502 | Naik |
| 9 | I | 4,630,276 | 3,811,740 | Turun |
| 10 | J | 16,235,397 | 17,611,842 | Naik |
| 11 | K | 2,810,321 | 3,357,161 | Naik |
| 12 | L | 9,949,829 | 14,110,933 | Naik |

Berikut merupakan visualisasi hasil perbandingan prediksi periode pertama dengan nilai aktual pada Gambar 2.



Gambar 2. Grafik Perbandingan Prediksi Periode Pertama dengan nilai Aktual

Pada Gambar 2. Terlihat Kecamatan yang mengalami kenaikan jumlah tunggakan pajak kendaraan jika dibandingkan dengan nilai aktual, di antaranya Kecamatan A, C, D, E, H, J, K dan Kecamatan L. Sementara Kecamatan B, F, G dan Kecamatan I mengalami penurunan jumlah tunggakan pajak.

3.5. Hasil Prediksi Terbaik untuk Periode Kedua (April–Juni 2019)

Untuk prediksi periode Kedua dilakukan langkah-langkah yang sama seperti pada periode Pertama, hasil prediksi terpilih periode Kedua diperoleh dari *fold-1* dengan nilai MSE = 0,0014 dan R= 0,9640. Model terbaik ini yang direkomendasikan sebagai hasil prediksi, dapat dilihat pada Tabel 5.

Hasil prediksi periode Kedua dibandingkan dengan nilai Aktual untuk melihat trend hasil prediksi. Pada tabel 7. Terdapat kenaikan jumlah tunggakan pajak di Kecamatan A, C, E, H, I, K dan Kecamatan L, dapat dilihat pada Tabel 6.

Tabel 5. Hasil Prediksi Periode Kedua (April-Juni) Tahun 2019 setelah di Denormalisasi.

| No | Nama Kecamatan | Prediksi Periode Kedua | Hasil Prediksi setelah di Denormalisasi (Rp) |
|----|----------------|------------------------|--|
| 1 | A | 0,242 | 22,345,343 |
| 2 | B | 0,369 | 32,739,833 |

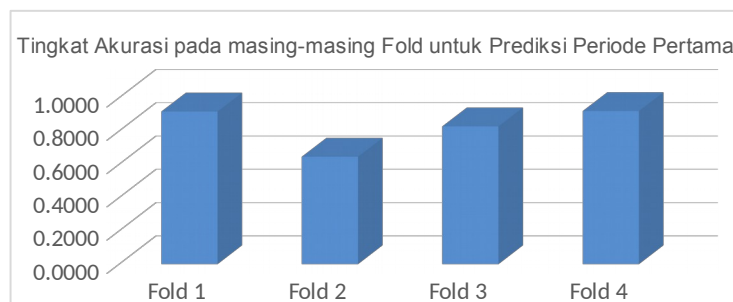
| No | Nama Kecamatan | Prediksi Periode Kedua | Hasil Prediksi setelah di Denormalisasi (Rp) |
|----|----------------|------------------------|--|
| 3 | C | 0,360 | 32,011,950 |
| 4 | D | 0,010 | 3,323,517 |
| 5 | E | 0,180 | 17,244,338 |
| 6 | F | 0,367 | 32,587,315 |
| 7 | G | 0,356 | 31,625,227 |
| 8 | H | 0,018 | 4,010,710 |
| 9 | I | 0,033 | 5,236,724 |
| 10 | J | 0,288 | 26,128,141 |
| 11 | K | 0,037 | 5,600,721 |
| 12 | L | 0,182 | 17,391,209 |

Tabel 6. Perbandingan hasil prediksi periode kedua dengan nilai aktual.

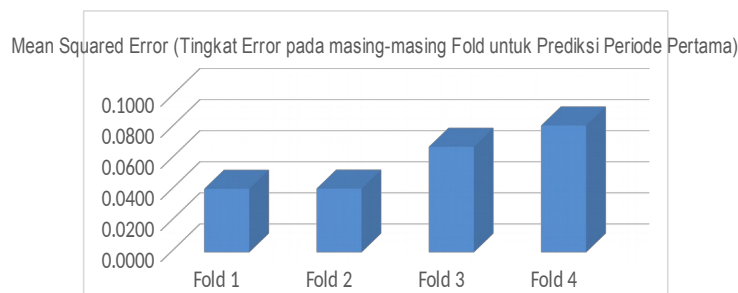
| No | Nama Kecamatan | Nilai Aktual | Hasil Prediksi Periode Kedua Tahun 2019 | Trend Hasil Prediksi |
|----|----------------|--------------|---|----------------------|
| 1 | A | 16,220,971 | 22,345,343 | Naik |
| 2 | B | 46,484,625 | 32,739,833 | Turun |
| 3 | C | 28,199,976 | 32,011,950 | Naik |
| 4 | D | 8,064,892 | 3,323,517 | Turun |
| 5 | E | 7,642,660 | 17,244,338 | Naik |
| 6 | F | 36,128,058 | 32,587,315 | Turun |
| 7 | G | 35,022,841 | 31,625,227 | Turun |
| 8 | H | 3,073,661 | 4,010,710 | Naik |
| 9 | I | 5,051,901 | 5,236,724 | Naik |
| 10 | J | 18,950,244 | 26,128,141 | Turun |
| 11 | K | 2,669,400 | 5,600,721 | Naik |
| 12 | L | 9,710,989 | 17,391,209 | Naik |

3.6. Analisis hasil prediksi menggunakan Algoritma *Support Vector Regression*

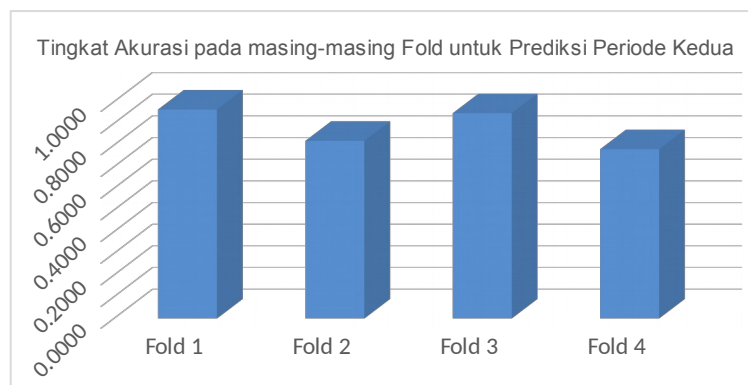
Berikut merupakan analisis hasil akurasi menggunakan algoritma *support vector regression*, di mana pengujian dengan 4 *fold* menghasilkan nilai MSE dan R untuk prediksi jumlah tunggakan pajak kedua periode. Nilai MSE mendekati 0 dan nilai R mendekati 1 membuktikan bahwa pengujian algoritma menghasilkan prediksi terbaik yang direkomendasikan sebagai hasil prediksi. Untuk hasil pengujian *fold* yang menghasilkan nilai MSE dan R pada Prediksi periode pertama dapat dilihat pada Gambar 3 dan 4, dan hasil akurasi untuk periode kedua dapat dilihat pada Gambar 5 dan 6.



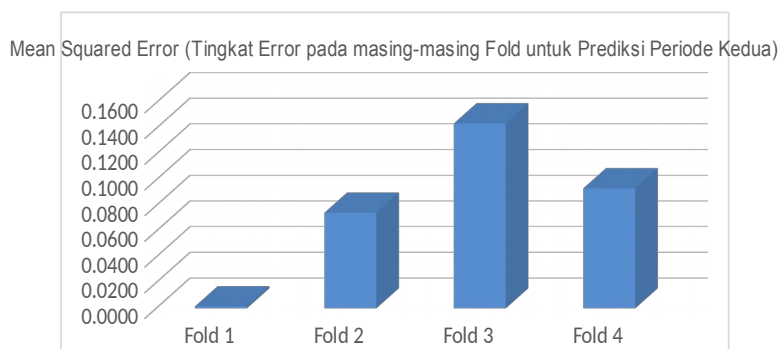
Gambar 3. Nilai R setiap *fold* untuk prediksi periode pertama.



Gambar 4. Nilai MSE setiap *fold* untuk prediksi periode pertama.



Gambar 5. Nilai R setiap *fold* untuk prediksi periode kedua



Gambar 6. Nilai MSE setiap *fold* untuk prediksi periode kedua.

4. Kesimpulan

Model terbaik yang didapatkan dari hasil prediksi jumlah tunggakan pajak kendaraan pada periode pertama terdapat pada *fold*-1 dengan akurasi 91% dan nilai error 4,8%. Untuk periode kedua model terbaik didapat dari *fold*-1 dengan akurasi 96% dan nilai error 0,14%. Hasil Prediksi Periode pertama menunjukkan sebanyak 8 Kecamatan mengalami kenaikan jumlah tunggakan pajak, dan 4 Kecamatan lainnya mengalami penurunan. Pada periode Kedua, sebanyak 7 Kecamatan mengalami peningkatan jumlah tunggakan pajak dan 5 Kecamatan lainnya mengalami penurunan tunggakan pajak.

Kenaikan jumlah tunggakan pajak salah satunya disebabkan oleh kepadatan penduduk yang meningkat di Kabupaten Pasaman, peningkatan diiringi dengan kebutuhan masyarakat akan sarana transportasi yakni Kendaraan. Peningkatan jumlah tunggakan pajak kendaraan akan berpengaruh terhadap pendapatan asli daerah yang nantinya mengganggu kelangsungan pembangunan daerah. Salah satu langkah yang direkomendasikan kepada Samsat adalah melakukan sosialisasi mengenai pentingnya membayar pajak tepat waktu karena faktor utama ketepatan membayar pajak adalah kesadaran dari wajib pajak sendiri, selanjutnya Samsat diharapkan meningkatkan pelayanan pajak guna memberikan kenyamanan wajib pajak dalam menunaikan kewajibannya.

Daftar Pustaka

- [1]. Ardianti, D. (2014). Pelaksanaan pemungutan pajak kendaraan bermotor oleh dinas pendapatan daerah provinsi jawa timur (studi kasus di kantor dinas pendapatan daerah provinsi jawa timur unit pelaksana teknis dinas (uptd) malang kota). *Kumpulan Jurnal Mahasiswa Fakultas Hukum*.
- [2]. Yuswanto, O. D. M. P. D., dkk. (2018). Pemanfaatan dana bagi hasil pajak kendaraan bermotor di kota bandar lampung. *JURNAL HIMA HAN*, 5(1).
- [3]. Siregar, Y. A., Saryadi, S., dan Listyorini, S. (2012). Pengaruh pelayanan fiskus dan pengetahuan perpajakan terhadap kepatuhan wajib pajak (studi empiris terhadap wajib pajak di Semarang tengah). *Jurnal Ilmu Administrasi Bisnis*, 1(2), 295–304.

- [4]. Prasad, M., Florence, L., dan Arya, A. (2015). A study on software metrics based software defect prediction using data mining and machine learning techniques. *International Journal of Database Theory and Application*, 8(3), 179–190.
- [5]. Hermawati, F. A. (2013). *Data Mining*. Yogyakarta: Penerbit Andi.
- [6]. Mustakim, M., Buono, A., dan Hermadi, I. (2015). Support vector regression untuk prediksi produktivitas kelapa sawit di provinsi riau. *Jurnal Sains dan Teknologi Industri*, 12(2), 179–188.
- [7]. Akande, K. O., Owolabi, T. O., Olatunji, S. O., dan Abdulraheem, A. (2016). A novel homogenous hybridization scheme for performance improvement of support vector machines regression in reservoir characterization. *Applied Computational Intelligence and Soft Computing*, 2016.
- [8]. Agustina, S. D., Bella, C., dan Ramadhan, M. A. (2018). Support vector regression algorithm modeling to predict the availability of foodstuff in indonesia to face the demographic bonus. Dalam *Journal of physics: Conference series* (Vol. 1028, hal. 012240).
- [9]. Milovic, B. (2012). Prediction and decision making in health care using data mining. *Kuwait chapter of arabian journal of business and management review*, 33(848), 1–11.
- [10]. David, J. (2007). Hand. principles of data mining. *Drug safety*, 30(7), 621–622.
- [11]. Maimon, O., dan Rokach, L. (2005). Introduction to supervised methods. Dalam *Data mining and knowledge discovery handbook* (hal. 149–164). Springer.
- [12]. Mustakim, M., Buono, A., dan Hermadi, I. (2016). Performance comparison between support vector regression and artificial neural network for prediction of oil palm production. *Jurnal Ilmu Komputer dan Informasi*, 9(1), 1–8.
- [13]. Smola, A. J., dan Scholkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3), 199–222.
- [14]. Kale, S., Kumar, R., dan Vassilvitskii, S. (2011). Cross-validation and mean-square stability. Dalam *In proceedings of the second symposium on innovations in computer science (ics2011)*.