

Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS)
dan Penerapannya pada Pemodelan Produk Domestik Regional Bruto (PDRB)
di Provinsi Sumatera Barat

Tika Mijayanti^{#1}, Helma^{*2}

[#]*Student of Mathematics Departement, Universitas Negeri Padang, Indonesia*

^{*}*Lecture of Mathematics Departement, Universitas Negeri Padang, Indonesia*

¹tika.mijayanti01@gmail.com

²helma_mat@fmipa.unp.ac.id

Abstract — Increased economic growth can help a region's economy grow and demonstrate that the government is capable of improving the welfare of its citizens. The rate of economic growth may be measured by Gross Regional Domestic Product (GRDP). This look at turned into performed to decide the factors that maximum effect GRDP inside the province of West Sumatera from 2015 to 2019 using *Bootstrap Aggregating Multivariate Adaptive Regression Splines* (Bagging MARS). The best model with the lowest GCV value is 7,36868 with BF=8, MI=3 and MO=0 as a combination. Then Bagging was carried out on the initial dataset with 50 Bootstrap replications to obtain the smallest GCV of 5,256292. Based on this, the smallest GCV value obtained from Bagging MARS is smaller than the MARS method. Meaning that the Bagging method can lessen the GCV value and increase accuracy. So that the factors that maximum influence GRDP in the province of West Sumatera are Regional Original Income.

Keywords — GRDP, GCV, MARS, Bagging

Abstrak — Peningkatan pertumbuhan ekonomi dapat memajukan perekonomian di suatu daerah serta menunjukkan bahwa pemerintah mampu meningkatkan kesejahteraan masyarakat di daerah. Laju pertumbuhan ekonomi bisa diukur dengan Produk Domestik Regional Bruto (PDRB). Penelitian ini dilakukan agar mengetahui faktor-faktor yang paling mempengaruhi PDRB di Provinsi Sumatera Barat dari tahun 2015 sampai tahun 2019 dengan metode *Bootstrap Aggregating Multivariate Adaptive Regression Splines* (Bagging MARS). Model terbaik yang dihasilkan dengan nilai *Generalized Cross Validation* (GCV) terkecil yaitu 7,36868 menggunakan kombinasi BF=8, MI=3 dan MO=0. Selanjutnya dilakukan Bagging pada dataset awal dengan 50 ulangan Bootstrap yang diperoleh GCV terkecil 5,256292. Berdasarkan hal tersebut, nilai GCV terkecil yang diperoleh dari Bagging MARS lebih kecil dibandingkan dengan metode MARS. Artinya metode Bagging bisa menurunkan nilai GCV serta meningkatkan akurasi. Oleh karena itu diperoleh faktor-faktor yang paling mempengaruhi PDRB di Provinsi Sumatera Barat adalah Pendapatan Asli Daerah.

Kata kunci — PDRB, GCV, MARS, Bagging

PENDAHULUAN

Rata-rata warga di negara maju mempunyai pendapatan 10 kali lebih tinggi daripada negara berkembang. Perbedaan pendapat yang signifikan ini mencerminkan perbedaan dalam kualitas hidup warga negaranya. Pada suatu negara, standar hidup akan berubah satu persatu sehingga menaikkan pertumbuhan ekonomi perlu dilakukan oleh negara-negara berkembang dalam mengejar ketinggalan di bidang ekonomi [1]. Sebagai negara berkembang, Provinsi di Indonesia perlu meningkatkan pertumbuhan ekonomi, agar perekonomian di suatu daerah mengalami kemajuan serta dapat meningkatkan kesejahteraan masyarakatnya. Peningkatan pertumbuhan ekonomi ditandai dengan meningkatnya

barang serta jasa yang diproduksi. Karena pertumbuhan ekonomi berkaitan dengan produksi barang serta jasa dalam aktivitas ekonomi warga, maka laju pertumbuhan ekonomi bisa diukur memakai data Produk Domestik Regional Bruto (PDRB).

PDRB ialah jumlah output bruto seluruh barang serta jasa pada suatu daerah yang diperoleh dari aktivitas ekonomi pada periode tertentu [2]. Nilai PDRB juga bisa menunjukkan kondisi dan pencapaian kinerja perekonomian di suatu daerah. Informasi ini sangat diperlukan sehingga mendukung setiap kebijakan yang akan diambil oleh pemerintahan pada suatu daerah. Pada penelitian ini digunakan nilai PDRB atas dasar harga konstan, yaitu penilaiannya berdasarkan di harga satu tahun tertentu.

TABEL I
LAJU PERTUMBUHAN PRODUK DOMESTIK REGIONAL BRUTO ATAS DASAR
HARGA KONSTAN 2010 MENURUT KABUPATEN/KOTA DI PROVINSI
SUMATERA BARAT (PERSEN), 2015-2019

Kabupaten/Kota	2015	2016	2017	2018	2019
Kep. Mentawai	5,20	5,02	5,12	4,89	4,73
Pesisir Selatan	5,73	5,33	5,41	5,32	4,78
Solok	5,44	5,31	5,32	5,19	5,04
Sijunjung	5,69	5,26	5,26	5,07	4,81
Tanah Datar	5,33	5,03	5,11	5,04	4,99
Padang Pariaman	6,14	5,52	5,58	5,44	2,40
Agam	5,52	5,51	5,43	5,23	4,78
Lima Puluh Kota	5,61	5,32	5,33	5,23	5,06
Pasaman	5,34	5,07	5,08	4,97	4,80
Solok Selatan	5,35	5,14	5,15	5,00	4,86
Dharmasraya	5,75	5,42	5,44	5,28	4,94
Pasaman Barat	5,70	5,34	5,34	5,21	4,45
Padang	6,41	6,17	6,23	6,06	5,65
Solok	5,97	5,76	5,76	5,65	5,49
Sawahlunto	6,03	5,72	5,74	5,50	5,31
Padang Panjang	5,91	5,80	5,80	5,71	5,56
Bukittinggi	6,14	6,05	6,08	6,00	5,84
Payakumbuh	6,19	6,08	6,12	6,02	5,89
Pariaman	5,79	5,59	5,61	5,47	5,30
Jumlah	5,53	5,27	5,30	5,14	5,01

Pada tabel I terlihat bahwa kinerja perekonomian Sumatera Barat tahun 2015 belum menunjukkan perkembangan yang memuaskan. Laju pertumbuhan ekonomi pada tahun 2015 tercatat 5,41 persen, pada tahun 2016 tercatat 5,27 persen yang sedikit rendah dari pada tahun sebelumnya, sedangkan tahun 2017 mengalami sedikit peningkatan menjadi 5,30 persen. Namun, pada tahun 2018 tercatat 5,14 persen dan tahun 2019 tercatat 5,01 persen kembali mengalami penurunan. Berdasarkan hal tersebut, pertumbuhan ekonomi harusnya mengalami kenaikan dari tahun ke tahun agar meningkatkan perubahan struktur perekonomian wilayah menuju perekonomian yang teratur dan bergerak maju. Sehingga perlu diketahui faktor-faktor yang paling mempengaruhi pertumbuhan ekonomi di Provinsi Sumatera Barat.

Beberapa faktor yang berpengaruh pada PDRB ialah Pendapatan Asli Daerah (PAD), Penanaman Modal Asing (PMA), Penanaman Modal Dalam Negeri (PMDN) serta Tingkat Partisipasi Angkatan Kerja (TPAK). PAD yaitu sumber penerimaan daerah yang berkaitan dengan pertumbuhan ekonomi. Selanjutnya, peningkatan penanaman modal baik dalam negeri maupun luar negeri bisa meningkatkan proses produksi barang maupun jasa sehingga dapat meningkatkan perekonomian serta meningkatkan standar hidup masyarakat suatu daerah. Berikutnya, tenaga kerja juga berperan dalam proses produksi yang menghasilkan barang sehingga meningkatkan tenaga kerja dapat meningkatkan pertumbuhan ekonomi. TPAK adalah perbandingan antara jumlah angkatan kerja terhadap penduduk padausia kerja. Nilai TPAK yang tinggi akan menunjukkan suatu daerah mempunyai kesempatan kerja yang besar sehingga dapat menaikkan pendapatan daerah dan sebaliknya nilai TPAK yang rendah menunjukkan rendahnya kesempatan kerja bagi penduduk dalam usia kerja di suatu daerah. Keberhasilan daerah dalam menunjukkan faktor-faktor dapat digunakan sebagai ukuran daya saing perekonomian antar daerah dalam meningkatkan daya tarik dan memnangkan persaingan ekonomi. Analisis regresi yang banyak digunakan adalah regresi parametrik yang diduga bentuk kurvanya (linier, kuadrat, kubik) sesuai teori yang bisa memberikan informasi hubungan [4].

Metode MARS merupakan suatu teknik nonparametrik yang dipakai untuk mendapati hubungan nonlinear antara variabel respon dan variabel prediktor. Karena MARS adalah metode regresi nonparametrik, maka model MARS tidak bergantung pada asumsi tertentu. Metode MARS juga bermanfaat dalam mengatasi permasalahan data yang berdimensi tinggi. Untuk menentukan model terbaik MARS, digunakan kombinasi nilai knot, basis fungsi, dan maksimum interaksi berdasarkan nilai *trial and error* untuk menemukan nilai GCV minimum. *Bootstrap Aggregating* (Bagging) pertama kali diperkenalkan oleh Breiman pada tahun 1996. Metode ini dapat dipakai untuk meningkatkan stabilitas, peningkatan akurasi, dan kekuatan prediksi. Bagging merupakan metode yang menggabungkan banyak nilai prediktor. Penggabungan banyak nilai dilakukan dengan penggandaan minimum sebanyak 50 kali [5].

Metode MARS adalah perluasan dari metode *Recursive Partitioning Regression* (RPR) yang dikombinasikan dengan metode *Spline*. Tetapi pendekatan dari RPR ini mempunyai kelemahan yaitu model yang didapatkan tidak kontinu pada knot. Menurut [6], pengembangan RPR yang dikombinasikan dengan pendekatan metode *Spline* menghasilkan model yang kontinu pada knot yang berarti garis regresi selalu menyambung serta setiap knot selalu menyambung dengan fungsi basisnya.

Setelah dilakukan pengembangan fungsi basis dengan model RPR menggunakan kombinasi *spline* maka diperoleh model MARS sebagai berikut:

$$f(x) = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{K_m} [S_{km}(x_{v(k,m)} - t_{km})]_+$$

Tingkat akurasi model MARS bisa ditingkatkan dengan metode Bagging. Sehingga, metode MARS akan dikombinasikan dengan metode Bagging yang dapat meningkatkan akurasi kekuatan prediktif model MARS.

METODE

Penelitian ini yaitu penelitian terapan serta menggunakan data sekunder yang didapatkan dari Badan Pusat Statistik yang dipublikasikan dalam bukunya termuat dalam [2-3]. Data tersebut mencakup 12 Kabupaten dan 7 Kota di Provinsi Sumatera Barat berupa data PDRB pada tahun 2015 sampai tahun 2019 sebagai variabel respon. Dan variabel prediktor adalah Pendapatan Asli Daerah (X_1), Penanaman Modal Asing (X_2), Penanaman Modal dalam Negeri (X_3) dan Tingkat Partisipasi Angkatan Kerja (X_4).

Langkah-langkah analisis data pada penelitian ini diantaranya:

1. Memperoleh data variabel respon serta variabel prediktor dan melakukan analisis deskriptif yang berguna untuk menunjukkan karakteristik data setiap variabel, kemudian melihat kemungkinan terdapatnya knot pada data dan melakukan uji asumsi klasik untuk melihat apakah dapat menggunakan pendekatan secara nonparametrik atau tidak.
2. Membagi data menjadi data *training* untuk membentuk model dan data *testing* untuk validasi data.
3. Memilih jumlah fungsi basis (BF) maksimum yang memberikan selang antar knot data. Maksimum fungsi basis yang diperbolehkan adalah sebesar 2-4 kali dari banyaknya peubah prediktor yang dipergunakan. Pada penelitian ini digunakan 4 peubah prediktor sehingga maksimum jumlah BF yaitu 8, 12, dan 16.
4. Menentukan jumlah maksimum Interaksi (MI). Maksimum interaksi yang diperbolehkan adalah 1, 2, dan 3. Jika diperoleh lebih 3 interaksi, akan menyebabkan penafsiran model yang kompleks.
5. Memilih nilai minimum observasi (MO) antar knot yaitu 0, 1, 2, dan 3.
6. Memperoleh model MARS terbaik untuk PDRB di Provinsi Sumatera Barat berdasarkan kriteria nilai GCV minimum yang dihasilkan dari kombinasi nilai BF, MI, dan MO.
7. Memperoleh variabel-variabel yang signifikan mempunyai pengaruh pada model MARS terbaik pada dataset awal.
8. Melakukan uji signifikansi model MARS untuk mengevaluasi kecocokan model menggunakan uji

serentak dan uji parsial.

9. Melakukan pemodelan Bagging MARS pada data menggunakan 50 kali replikasi.
10. Menentukan model terbaik metode Bagging MARS.
11. Menentukan variabel terbaik yang paling mempengaruhi.

HASIL DAN PEMBAHASAN

A. Analisis Deskriptif

Analisis deskriptif dilakukan agar mengetahui karakteristik pertumbuhan Kabupaten/Kota pada Provinsi Sumatera Barat tahun 2015 sampai 2019 menggunakan metode *Bootstrap Aggregating Multivariate Adaptive Regression Splines* Bagging MARS.

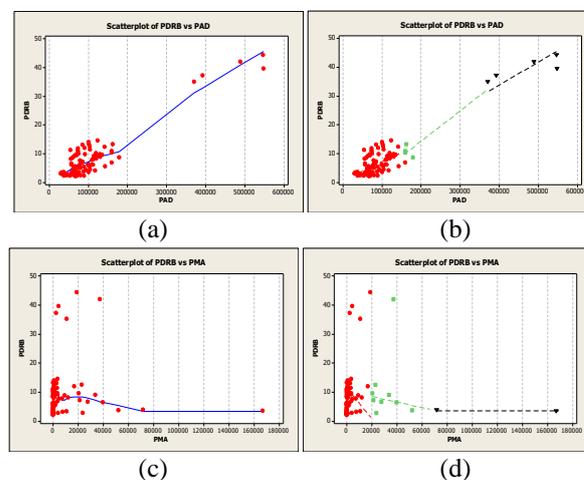
TABEL III
ANALISIS DESKRIPTIF DATA PENELITIAN

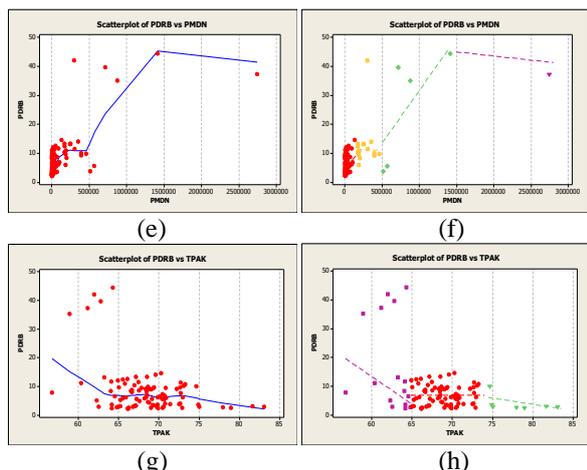
Peubah	Rata-rata	Standar Deviasi	Min	Maks
Y	8,286708	8,190256	2,006625	4,44593
X ₁	103817,41	93938,927	29692,04	547764,96
X ₂	7039,36	20306,183	0,00	167093,2
X ₃	128432,88	342051,21	0,00	2745716,4
X ₄	68,44	4,299953	56,74	83,1

Berdasarkan tabel II diperoleh bahwa standar deviasi untuk variabel PDRB (Y), X₁ dan X₄ relatif rendah yang menunjukkan sebaran data pada variabel yang tertera menyebar disekitar rata-rata atau dapat diartikan tidak terdapat perbedaan yang besar pada sebaran data masing-masing variabel. Variabel X₂ dan X₃ memiliki nilai standar deviasi relatif besar yang menunjukkan sebaran data tersebut menunjukkan sebaran data menyebar cukup jauh dari nilai rata-rata.

1) Knot pada data

Pada langkah ini akan dilihat apakah terdapat knot pada data dengan melihat *Smoothing Lines* atau dengan melakukan pembagian data variabel prediktor menjadi beberapa segmen membentuk sifat *Piecewise* seperti plot data berikut:





Gambar 1. Smoother Lines dan Grafik Perubahan Garis Regresi Data

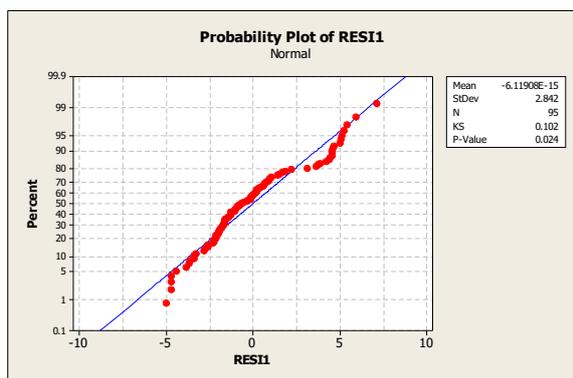
Gambar 1 adalah *Scatterplot Smoother Lines* dan perubahan garis regresi yang tersegmentasi pada nilai PDRB (Y) terhadap masing-masing variabel prediktor. Pada gambar 1 diperoleh pola data yang menyebar dan tidak memperlihatkan kecenderungan membentuk pola tertentu, oleh karena itu dapat diduga terdapat knot pada data yang dijelaskan pada masing-masing *Smoother Lines* nilai PDRB (Y) terhadap masing-masing variabel prediktor pada bagian (a), (c), (e), dan (g) yang tidak membentuk garis lurus. Bagian (b), (d), (f), dan (h) merupakan perubahan garis regresi yang tersegmentasi membentuk garis regresi yang tidak saling terhubung mengikuti suatu pola dan data berada pada *slope* yang berbeda.

2) Uji Asumsi Klasik

Langkah berikutnya melakukan pengujian pada asumsi regresi untuk data PDRD terhadap data faktor-faktor yang diduga mempunyai pengaruh.

a. Uji Normalitas

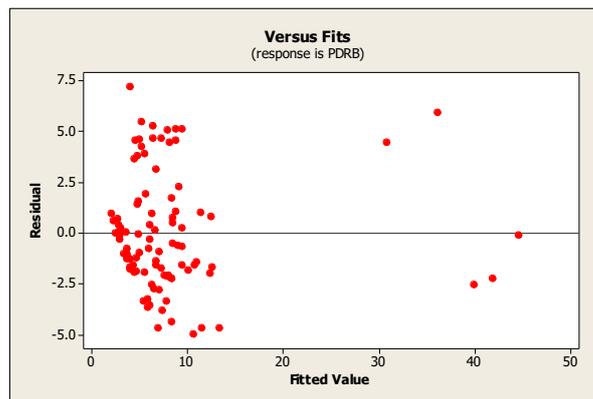
Uji normalitas memiliki tujuan untuk melihat apakah variabel berdistribusi normal atau tidak. Pengujian residual dengan metode Kolmogorov smirnov.



Gambar 2. Uji Normalitas

Pada gambar 2 menghasilkan nilai $p - value < \alpha$ atau 0,024 lebih rendah dari taraf signifikansi $\alpha(0,05)$ sehingga tolak H_0 yang artinya galat tidak berdistribusi normal.

b. Uji Homoskedastisitas



Gambar 3. Uji Homoskedastisitas

Pada gambar 3 terlihat bahwa titik-titik pada grafik menyebar atau tidak membuat pola tertentu sehingga asumsi homoskedastisitas terpenuhi atau dapat disimpulkan tidak terjadi heterokedastisitas.

c. Uji Autokorelasi

Langkah berikutnya untuk memeriksa adanya korelasi pada data pengamatan dapat dilakukan dengan Uji Durbin-Watson. Dengan banyak pengamatan (N)=95 dan banyak variabel bebas (k)=4, diperoleh hasil pengujian 1,28999. Pada tabel Durbin Watson dengan taraf signifikansi 0,05 diperoleh nilai $d_U=1,7546$ dan $d_L=1,5795$. Karena nilai $d_{hitung} < d_U$ maka dapat disimpulkan bahwa tolak H_0 atau terdapat korelasi.

d. Uji Multikolinearitas

Adanya multikolinearitas dapat dilihat dari nilai *Variance Inflation Faktor* (VIF). Perolehan nilai VIF untuk variabel prediktor $X_1=1,993$, $X_2=1,051$, $X_3=1,824$ dan $X_4=1,213$. Nilai variabel prediktor $X_1, X_2, X_3, X_4 \leq 10$ maka bisa disimpulkan tidak terjadi multikolinearitas antar variabel prediktor atau tidak terdapat korelasi antar variabel tersebut.

B. Pembentukan Model PDRB Menggunakan Metode MARS

Berdasarkan plot data diketahui adanya keberadaan knot dan hasil pengujian asumsi klasik yang dilakukan adalah uji normalitas, uji homoskedastisitas, uji autokorelasi, dan uji multikolinearitas didapatkan asumsi yang tidak terpenuhi yaitu galat yang tidak berdistribusi normal dan terjadinya autokorelasi. Oleh karena itu, pendekatan statistik nonparametrik dapat digunakan dalam penelitian ini untuk memperoleh hasil model terbaik dengan menggunakan metode MARS.

Langkah awal yang dilakukan adalah membagi data *training* serta data *testing*. Dimana data *training* untuk membentuk model dan data *testing* untuk memvalidasi data. Agar analisis selanjutnya menggunakan perbandingan data yang tepat, hasil perbandingan data ditunjukkan pada tabel III:

TABEL III
PERBANDINGAN DATA TRAINING DAN DATA TESTING

Perbandingan Data	GCV	R ²	MSE
Training 70% Testing 30%	7,39672	0,929	4,03739
Training 75% Testing 25%	7,59997	0,922	4,31522
Training 80% Testing 20%	7,75293	0,903	4,99457
Training 85% Testing 15%	7,36868	0,892	5,26988

Pada tabel III diperoleh perbandingan data *training* serta data *testing* yang tepat dengan melihat nilai GCV terkecil. Jika terdapat nilai GCV yang sama, maka bisa dilihat dari nilai MSE terkecil. Jika terdapat nilai MSE yang sama, maka dapat dilihat dari R² terbesar. Nilai GCV terkecil terletak pada perbandingan data *training* (70%) sebanyak 81 data dan data *testing* (30%) sebanyak 14 data.

Selanjutnya, untuk menentukan model MARS didapatkan berdasarkan hasil kombinasi antara jumlah maksimum Basis Fungsi (BF), jumlah Maksimum Interaksi (MI) dan jumlah Maksimum Observasi (MO) dengan cara *trial and error* yang diinputkan dalam MARS menggunakan *software* SPM 8.0. Banyak variabel prediktor yang digunakan pada penelitian ini adalah 4 variabel oleh karena itu banyaknya basis fungsi yang digunakan adalah 8, 12 dan 16. Berikut *trial and error* nilai GCV, R², dan MSE dengan kombinasi BF, MI, MO terhadap data penelitian:

Model terbaik yang diperoleh dengan kombinasi BF=8, MI=3, MO=0 menghasilkan nilai GCV minimum yaitu 7,36868, nilai R² 0,892 dan nilai MSE 5,26988. Model PDRB di Provinsi Sumatera Barat berdasarkan Kabupaten/Kota pada tahun 2015 sampai tahun 2019 diantaranya:

$$Y = 6,30876 + (7,26827 \cdot 10^{-5}) BF1 - (2,09376 \cdot 10^{-10}) BF3 - (0,00379902) BF5 - (1,6505 \cdot 10^{-10}) BF7;$$

dengan basis fungsi (BF) dalam model sebagai berikut:

$$BF1 = \max(0, PAD - 29692);$$

$$BF3 = \max(0, 251928 - PMDN) BF1;$$

$$BF5 = \max(0, 746 - PMA);$$

$$BF7 = \max(0, TPAK - 72.97) BF3;$$

Berdasarkan model terbaik yang diperoleh, dapat ditarik kesimpulan bahwa peubah yang berpengaruh pada model yaitu peubah PAD (X₁) dengan tingkat kepentingan peubah 100 persen, peubah PMA (X₂) dengan tingkat kepentingan 18,57 persen, peubah PMDN (X₃) dengan tingkat kepentingan 23,26 persen dan TPAK (X₄) dengan tingkat kepentingan 6,63 persen. Dapat dikatakan bahwa semua peubah prediktor berpengaruh terhadap perkembangan PDRB di Provinsi Sumatera Barat dengan peubah prediktor yang paling

mempengaruhi PDRB di Provinsi Sumatera Barat yaitu Pendapatan Asli Daerah.

C. Uji Signifikansi Model

Hasil uji signifikansi model PDRB di Provinsi Sumatera Barat untuk memeriksa signifikansi parameter model terbaik sebagai berikut:

1) Uji serentak

Hipotesis uji serentak dilakukan sebagai berikut:

$$H_0 : a_1 = a_3 = a_5 = a_7$$

$$H_1 : \text{ada } a_m \neq 0$$

Dengan taraf signifikansi $\alpha = 0,05$

TABEL IV
UJI SERENTAK PADA MODEL

F-Statistik	157,37492
P-Value	0,00000
[MDF,NDF]	[4, 76]
Standar Error	2,36993
Residual Sum Of Squares	426,86015
Regression Sum Of Squares	3535,63598

Berdasarkan tabel IV, uji serentak pada parameter model terbaik didapatkan $p - value = 0,0000$, F-Statistik=157,37492 serta nilai $F_{0,05(4;76)} = 2,4$ berdasarkan tabel F. Sehingga diperoleh $p - value < \alpha$ atau $F_{hitung} > F_{0,05(4;76)}$ dan dapat disimpulkan bahwa H_0 ditolak yang berarti model signifikan dan dapat digunakan dalam mengukur perkembangan PDRB di Sumatera Barat.

2) Uji Parsial

Hipotesis uji parsial dilakukan sebagai berikut:

$$H_0 = a_m = 0$$

$$H_1 = a_m \neq 0$$

$m = 1, 3, 5$ dan 7

Dengan taraf signifikan $\alpha = 0,05$

TABEL V
UJI PARSIAL PADA MODEL

Peubah	Dugaan	Standar Error	T-Rasio	P-Value
Konstan	6,30876	0,63416	9,94826	0,00000
Basis Fungsi 1	0,00007	0,00000	21,12663	0,00000
Basis Fungsi 3	-0,00000	0,00000	-5,81430	0,00000
Basis Fungsi 5	-0,00380	0,00082	-4,62186	0,00002
Basis Fungsi 7	-0,00000	0,00000	-2,89968	0,00488

Berdasarkan tabel V, hasil pengujian parsial pada parameter terbaik untuk masing-masing BF yang masuk pada model memiliki nilai $p - value < \alpha$. Oleh karena itu bisa ditarik kesimpulan bahwa H_0 ditolak yang artinya koefisien a_m untuk $m=1, 3, 5$ dan 7 berpengaruh terhadap model.

D. Interpretasi Model

1. $BF1 = \max(0, PAD - 29692)$ dengan koefisien $7,26827 \cdot 10^{-5}$ yang berarti bahwa setiap kenaikan BF1 akan meningkatkan nilai PDRB di Sumatera Barat (Y) sebesar $7,26827 \cdot 10^{-5}$ jika nilai Pendapatan Asli Daerah memperoleh nilai yang lebih besar dari 29692.
2. $BF3 = \max(0, 251928 - PMDN)$ BF1 dengan koefisien $-2,09376 \cdot 10^{-10}$ yang berarti bahwa setiap kenaikan BF3 akan menurunkan nilai PDRB di Sumatera Barat (Y) sebesar $2,09376 \cdot 10^{-10}$ jika nilai Penanaman Modal Dalam Negeri memperoleh nilai yang lebih kecil dari 251928 dan nilai nilai Pendapatan Asli Daerah memperoleh nilai yang lebih besar dari 29692.
3. $BF5 = \max(0, 746 - PMA)$ dengan koefisien $-0,00379902$ yang berarti bahwa setiap kenaikan BF5 akan menurunkan nilai PDRB di Sumatera Barat (Y) sebesar $0,00379902$ jika nilai Penanaman Modal Asing memperoleh nilai yang lebih kecil dari 746.
4. $BF7 = \max(0, TPAK - 72,97)$ BF3 dengan koefisien $-1,6505 \cdot 10^{-10}$ yang berarti bahwa setiap kenaikan BF7 akan menurunkan nilai PDRB di Provinsi Sumatera Barat (Y) sebesar $1,6505 \cdot 10^{-10}$ jika nilai Tingkat Partisipasi Angkatan Kerja memperoleh nilai yang lebih besar dari 72,97, nilai Penanaman Modal Dalam Negeri memperoleh nilai yang lebih kecil dari 251928 dan nilai nilai Pendapatan Asli Daerah memperoleh nilai yang lebih besar dari 29692.

E. Menentukan Model Bagging MARS

Setelah didapatkan model MARS terbaik, selanjutnya dilakukan pendekatan Bagging untuk mengetahui perbandingan hasil terbaik antara model MARS dengan model Bagging MARS. Pendekatan Bagging MARS dilakukan dengan melakukan replikasi sebanyak 50 kali dan didapatkan model terbaik pada replikasi ke-33 menggunakan nilai GCV terkecil. Hasil replikasi Bagging MARS seraca lengkap dapat diperoleh dari olah data *software* R. Pada model terbaik Bagging MARS, didapatkan variabel prediktor yang menunjukkan angka signifikan, yaitu variabel X_1, X_2, X_3 dengan nilai GCV sebesar 5,256292 dan nilai RSq 0,913606 dengan model:

$$\hat{f}(x) = 52,37468 + 0,0001091534 \max(PAD - 142764) - 0,006584142 \max(PMA - 772,6) - 0,006226118 \max(7424,3 - PMA) + 0,007021776 \max(PMA - 7424,3) - 0,0005605113 \max(PMA - 16595,7) - 1,492095 \cdot 10^{-5} \max(173892 - PMDN)$$

Berdasarkan hasil yang didapatkan dari model MARS dan Bagging MARS, maka:

TABEL VI
PERBANDINGAN MODEL MARS DENGAN MODEL BAGGING MARS

	Variabel Signifikan	Nilai GCV
MARS	X_1, X_2, X_3, X_4	7,36868
Bagging MARS	X_1, X_2, X_3	5,256292

Pada tabel VI dapat diketahui bahwa pada model MARS memperoleh nilai GCV sebesar 7,36868 dan nilai GCV di model Bagging MARS sebesar 5,256292. Berdasarkan perbandingan kedua model tersebut, nilai GCV pada model Bagging MARS menunjukkan angka yang lebih rendah. Hal ini menjelaskan bahwa Bagging dapat menurunkan nilai GCV dan menaikkan akurasi yang lebih baik dibandingkan menggunakan model MARS.

SIMPULAN

Bagging merupakan metode yang bisa digunakan pada metode klasifikasi serta regresi untuk mereduksi variansi pada suatu prediktor dan bisa menyempurnakan proses pendugaan serta berguna untuk meningkatkan akurasi pada data yang berdimensi tinggi sehingga Bagging bisa menurunkan nilai GCV dan menaikkan akurasi untuk metode MARS.

Model terbaik untuk PDRB di Provinsi Sumatera Barat menurut Kabupaten/Kota di tahun 2015 sampai tahun 2019 dengan metode *Bootstrap Aggregating Multivariate Adaptive Regression Splines* (Bagging MARS) diperoleh dengan nilai GCV terkecil sebesar 5,256292 dengan R^2 sebesar 0,913606 dan bentuk persamaan model sebagai berikut:

$$\hat{f}(x) = 52,37468 + 0,0001091534 \max(PAD - 142764) - 0,006584142 \max(PMA - 772,6) - 0,006226118 \max(7424,3 - PMA) + 0,007021776 \max(PMA - 7424,3) - 0,0005605113 \max(PMA - 16595,7) - 1,492095 \cdot 10^{-5} \max(173892 - PMDN)$$

Faktor-faktor yang berdampak signifikan PDRB menurut Kabupaten/Kota di Provinsi Sumatera Barat tahun 2015 sampai tahun 2019 menggunakan model Bagging MARS yaitu Pendapatan Asli Daerah (X_1), Penanaman Modal Asing (X_2), Penanaman Modal Dalam Negeri (X_4) dan variabel yang paling berpengaruh yaitu Pendapatan Asli Daerah (X_1).

REFERENSI

- [1] Mankiw, N. Gregory. 2001. *Pengantar Ekonomi Jilid 2*. Jakarta : Erlangga.
- [2] Badan Pusat Statistik. 2021. Provinsi Sumatera Barat dalam Angka 2021. Padang: Badan Pusat Statistik
- [3] Badan Pusat Statistik. 2019. Provinsi Sumatera Barat dalam Angka 2021. Padang: Badan Pusat Statistik
- [4] Draper, N.R., dan Smith, H. 1992. *Analisis Regresi Terapan*. Jakarta: PT. Gramedia Utama.
- [5] Breiman, L. 1996. Bagging Predistors. *Jurnal Machine Learning*. 24: 123-140
- [6] Friedman, Jerome H. 1991. Multivariate Adaptive Regression Splines. *The Annals of Statistic*, 19(1)