

Prediksi Keikutsertaan Pelaku Usaha dalam Pemanfaatan Insentif Pajak dengan *Artificial Neural Network*

Arifin Rosid

Direktorat Jenderal Pajak

Galih Ardin

Direktorat Jenderal Pajak

Tri Bayu Sanjaya

Direktorat Jenderal Pajak

Hak cipta © 2021 oleh Arifin Rosid, Galih Ardin, & Tri Bayu Sanjaya

Tulisan ini tidak mewakili pandangan Direktorat Jenderal Pajak (DJP).
Pandangan yang diungkapkan dalam tulisan ini adalah pendapat pribadi penulis dan tidak mencerminkan posisi resmi atau kebijakan dari DJP.

Prediksi Keikutsertaan Pelaku Usaha dalam Pemanfaatan Insentif Pajak dengan *Artificial Neural Network*

Arifin Rosid^{ID}
Direktorat Jenderal Pajak

Galih Ardin
Direktorat Jenderal Pajak

Tri Bayu Sanjaya
Direktorat Jenderal Pajak

Hak Cipta © 2021 oleh Arifin Rosid, Galih Ardin, dan Tri Bayu Sanjaya

Working paper ini dalam bentuk draf; didistribusikan dengan tujuan mendapatkan masukan atau bahan diskusi. Dilarang memproduksi ulang tanpa ijin penulis.

Prediksi Keikutsertaan Pelaku Usaha dalam Pemanfaatan Insentif Pajak dengan *Artificial Neural Network*

Arifin Rosid^{a,*}, Galih Ardin^a, & Tri Bayu Sanjaya^a

^aDirektorat Jenderal Pajak
Kementerian Keuangan Republik Indonesia

Abstract

Aiming at cushioning the economic impact and preserving the business capacity, one of key fiscal policies undertaken by the Indonesian government during the Covid-19 pandemic is the provision of tax incentives. Certain characteristics of businesses related to the use of these incentives are important information for policy makers. Unfortunately, there are no previous empirical studies that specifically identify the certain characteristics of Indonesian businesses that have a predictive function. This study offers an Artificial Neural Network (ANN) approach to predict the participation of businesses in tax incentives based on their main characteristics. The ANN model in this study utilises empirical data related to the number of labours, primary market share, annual turnover, nature of the main business activities, and main source of supply from surveyed 12,361 businesses. In this study, the ANN approach predicts with an accuracy rate almost 70%. The results indicate that the number of employees, annual turnover, and the primary market share of the businesses are the three most important variables that determine the participation of Indonesian businesses on utilising tax incentives.

Abstrak

Bertujuan untuk meredam dampak ekonomi dan menjaga kapasitas bisnis, salah satu kebijakan fiskal penting yang diambil oleh pemerintah Indonesia dalam masa pandemi Covid-19 adalah pemberian insentif pajak. Karakteristik tertentu dari pelaku usaha yang terkait dengan pemanfaatan insentif pajak merupakan informasi penting untuk diketahui oleh pengambil kebijakan. Sayangnya, belum ada studi empiris yang secara spesifik mengidentifikasi karakteristik pelaku usaha yang memiliki fungsi prediksi di Indonesia. Studi ini menawarkan pendekatan *Artificial Neural Network* (ANN) untuk memprediksi keikutsertaan pelaku usaha dalam insentif pajak berdasarkan karakteristik yang dimiliki. Model ANN dalam studi ini menggunakan data empiris jumlah pekerja, pangsa pasar utama, besaran omzet tahunan, sifat usaha utama, dan sumber utama pasokan dari 12.361 pelaku usaha hasil survei. Pendekatan ANN dalam studi ini memprediksi dengan tingkat akurasi hampir 70%. Hasil studi ini menunjukkan bahwa jumlah pekerja, omzet tahunan, dan pangsa pasar utama adalah tiga variabel terpenting yang menentukan keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak.

Kata kunci: Indonesia, Covid-19, insentif pajak, *Artificial Neural Network*

Kode klasifikasi JEL: D04, H20, H25, H30

* Alamat korespondensi penulis: Direktorat Potensi, Kepatuhan, dan Penerimaan, Kantor Pusat Direktorat Jenderal Pajak, Jalan Gatot Subroto No 40-42, Kebayoran Baru, Jakarta Selatan 12190. E-mail: arifin.rosid@gmail.com

1 PENDAHULUAN

1.1 Latar Belakang

Badan Pusat Statistik (2020a)–yang selanjutnya disingkat BPS–mencatat bahwa selama tahun 2020 terjadi penurunan aktivitas ekonomi yang disebabkan oleh pandemi Covid-19. Penurunan aktivitas ini tercermin dari anjloknya pertumbuhan ekonomi pada triwulan II sampai dengan triwulan IV tahun 2020 yang terjadi pada hampir semua sektor ekonomi. Lebih lanjut, berdasarkan survei yang dilakukan oleh BPS (2020b) juga diketahui bahwa pandemi Covid-19 telah menyebabkan penurunan omzet, kenaikan biaya, dan peningkatan angka pemutusan hubungan kerja pada tahun 2020.

Banyak negara telah mengambil kebijakan yang tegas untuk menahan dan mengurangi dampak ekonomi dari pandemi Covid-19 (OECD 2020). Guna mengurangi dampak negatif yang diakibatkan oleh pandemi Covid-19 khususnya pada sektor ekonomi, pemerintah Indonesia juga telah menerbitkan berbagai kebijakan ekonomi salah satunya pemberian insentif perpajakan. Insentif perpajakan diperlukan sebagai peredam awal untuk guncangan ekonomi yang disebabkan oleh anjloknya permintaan dan penawaran agregat.¹ Selain itu, insentif perpajakan juga diperlukan untuk membantu Wajib Pajak terdampak pandemi Covid-19 dan membantu pemulihan ekonomi nasional.

¹ Konsepsi skema kebijakan selama dan pasca pandemi menurut OECD (2020) dibagi dalam tiga fase: (i) respon yang segera; ii) meredam dampak ekonomi dan mempertahankan kapasitas usaha; dan (iii) pemulihan. Salah satu fokus kebijakan pajak dalam fase pertama dan kedua adalah sisi likuiditas (OECD 2020). Secara umum, insentif pajak yang diberikan oleh pemerintah melalui Peraturan Menteri Keuangan Nomor PMK-23/PMK.03/2020 sebagaimana telah diubah terakhir dengan PMK-82/PMK.03/2021 berkaitan erat dengan upaya membantu sisi likuiditas pelaku usaha.

Setidaknya, ada lima jenis insentif perpajakan yang diberikan pemerintah bagi Wajib Pajak terdampak pandemi Covid-19 sebagaimana diatur dalam Peraturan Menteri Keuangan Nomor PMK-23/PMK.03/2020 sebagaimana telah diubah terakhir dengan PMK-82/PMK.03/2021. Bentuk insentif pajak yang diberikan di antaranya adalah (i) insentif PPh Pasal 21 ditanggung Pemerintah (DTP), (ii) insentif pembebasan PPh Pasal 22 impor, (iii) insentif PPh Final peredaran bruto tertentu DTP, (iv) insentif pengurangan angsuran PPh Pasal 25, dan (v) insentif pengembalian pendahuluan PPN dipercepat. Berdasarkan Laporan Keuangan Pemerintah Pusat (LKPP) Tahun 2020 (*audited*) (Kementerian Keuangan RI 2021) diketahui bahwa sampai dengan akhir Desember tahun 2020, dari keseluruhan Wajib Pajak badan dan orang pribadi yang memenuhi kriteria, jumlah Wajib Pajak yang mengajukan permohonan insentif perpajakan adalah baru sejumlah 495.817 Wajib Pajak, dengan rincian: 146.068 Wajib Pajak mengajukan permohonan insentif PPh Pasal 21 DTP, 18.753 Wajib Pajak mengajukan permohonan insentif pembebasan PPh Pasal 22 impor, 79.796 Wajib Pajak mengajukan permohonan insentif pengurangan angsuran PPh Pasal 25, 2.593 Wajib Pajak mengajukan permohonan insentif pengembalian pendahuluan PPN dipercepat, dan 248.607 Wajib Pajak mengajukan permohonan insentif PPh Final peredaran bruto tertentu DTP. Total nilai permohonan insentif pajak selama tahun 2020 adalah sebesar Rp 46,11 triliun (Kementerian Keuangan RI 2021). Membandingkan jumlah Wajib Pajak

yang memanfaatkan insentif dengan jumlah Wajib Pajak yang ada, besar kemungkinan tingkat keikutsertaan para Wajib Pajak atau para pelaku usaha belum maksimal.²

Terkait hal ini, OECD (2020) merekomendasikan agar insentif pajak ditujukan kepada pelaku usaha yang paling membutuhkan dukungan. Meskipun secara administratif tidak mudah, pemberian stimulus ke target yang tepat akan membantu proses pemulihan bagi pelaku usaha yang paling membutuhkan dengan lebih cepat. Dengan demikian, kemampuan pemerintah untuk mengidentifikasi dan mendorong pelaku usaha agar mendapatkan insentif yang paling diperlukan menjadi prasyarat penting.

Sayangnya, mengidentifikasi faktor-faktor yang relevan dengan keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak bukan pekerjaan yang mudah. Selain membutuhkan ketersediaan data empiris yang memadai, diperlukan juga metode analisis data yang tepat. Selain itu, faktor-faktor yang berkaitan dengan keikutsertaan pelaku usaha dalam memanfaatkan insentif pajak banyak berkaitan dengan faktor-faktor yang bersifat non-linier, misalnya pangsa pasar, sumber utama pasokan, sifat usaha, atau lokasi usaha. Diantara kelebihan dari *neural network*, selain bersifat adaptif terhadap data, adalah kapabilitasnya dalam *non-linieritas* (Haykin 1999). Dalam area yang bersifat non-linier, tingkat akurasi dari metode yang ada dalam kecerdasan buatan (*artificial intelligence*) cenderung lebih unggul dari pada metode statistik tradisional (Bahrammirzaee 2010).

² Sebagai gambaran, pada tahun 2019 terdapat 3.323.971 pelaku usaha berbentuk Wajib Pajak badan (DJP 2021). Dalam tulisan ini penggunaan istilah pelaku usaha lebih sering digunakan untuk menggantikan istilah Wajib Pajak.

Beranjak dari sini, tim penulis mencoba menggunakan teknis analisis yang relevan untuk memprediksi dan mengidentifikasi karakteristik atau faktor-faktor yang menentukan tingkat keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak. Salah satu metode analisis data yang populer digunakan dalam area kecerdasan buatan adalah *Artificial Neural Network* (ANN). ANN adalah alat pemodelan non-parametrik yang dapat melakukan pemetaan fungsi-fungsi yang kompleks dengan tingkat akurasi yang memadai (Zhang et al. 1999).

ANN cukup banyak digunakan dalam penelitian lintas disiplin ilmu. Misalnya, di bidang keuangan, Sánchez-Serrano et al. (2020) menggunakan pendekatan ANN untuk membuat model prediksi dari opini audit khusus untuk laporan keuangan terkonsolidasi. Studi ini dapat memprediksi opini audit dengan tingkat akurasi 83%. Penggunaan pendekatan *artificial intelligence* di bidang keuangan mulai banyak digunakan mengingat seringkali perilaku bersifat non-linier dan penuh ketidakpastian (Bahrammirzaee 2010). Di bidang pendidikan misalnya, Aryadoust dan Baghaei (2016) menguji hubungan antara kemampuan membaca, pengetahuan leksikal, dan tata bahasa dari kelompok siswa yang menggunakan bahasa Inggris sebagai bahasa asing. Dalam studi ini, ANN dapat secara akurat mengklasifikasikan secara akurat sekitar 78% siswa.

Pendekatan ANN juga digunakan dalam studi perpajakan (lihat misalnya, Denton et al. 1995; Chen et al. 2011; Lin et al. 2012; Jupri & Sarno 2018; Jang 2019; Pérez López et al. 2019). Misalnya, pendekatan ANN diaplikasikan untuk melakukan proyeksi besarnya penerimaan pajak di Korea Selatan (Jang 2019). Lin et al. (2012) menggunakan

pendekatan *artificial intelligence* untuk mendeteksi penyelundupan pajak (*tax evasion*) di Taiwan. Meski demikian, sepanjang pengetahuan penulis, belum ada studi empiris menggunakan pendekatan ANN untuk memprediksi keikutsertaan pelaku usaha dalam insentif pajak, khususnya yang diberikan selama pandemi Covid-19.

Tulisan ini berkontribusi di dua tataran. Pertama, pada tataran literatur perpajakan, tulisan ini berkontribusi dalam memberikan pengetahuan tambahan mengenai bagaimana pendekatan *artificial intelligence* dapat diterapkan dalam area studi perpajakan. Implikasi teoritis maupun empiris dari studi ini dapat dibilang cukup signifikan karena merupakan studi yang pertama kali mencoba memprediksi keikutsertaan pelaku usaha dalam memanfaatkan insentif pajak selama pandemi menggunakan data observasi level nasional dengan jumlah cukup besar. Kedua, dalam tataran praktis, hasil analisis ini menghasilkan informasi yang relevan bagi pengambil kebijakan, baik dalam fase desain maupun implementasi kebijakan, khususnya terkait dengan pemanfaatan insentif pajak.

Tulisan ini terdiri dari empat bagian utama. Setelah membahas tujuan studi, tulisan ini akan membahas mengenai metode penelitian. Bagian ini meliputi data yang digunakan dalam studi dan pendekatan yang digunakan untuk menganalisis data. Selanjutnya, akan diuraikan mengenai hasil analisis dan pembahasan dari hasil analisis. Simpulan dan saran menjadi bagian penutup dari tulisan ini.

1.2 Tujuan Studi

Studi empiris ini bertujuan untuk menjawab tiga pertanyaan riset. Pertama, seberapa akurat pendekatan ANN dapat memprediksi status keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak? Kedua, tiga faktor apa yang paling relevan dalam memprediksi keikutsertaan pelaku usaha dalam memanfaatkan insentif pajak?³ Ketiga, seberapa kokoh (*robust*) hasil prediksi ANN jika sampel data yang digunakan dilakukan *sampling variability* berbasis lokasi pelaku usaha? Pertanyaan ketiga ini sekaligus bertujuan untuk mengetahui apakah karakteristik utama dari pelaku usaha yang teridentifikasi memiliki hubungan kuat dengan keikutsertaan pelaku usaha dalam insentif pajak di enam kelompok pulau (Jawa, Sumatra, Kalimantan, Sulawesi, Bali – Nusa Tenggara, dan Papua – Maluku) bersifat homogen.

2 METODE PENELITIAN

2.1 Data

Studi ini menguji data dari 12.361 pelaku usaha yang dikumpulkan melalui survei.⁴ Pelaku usaha yang menjadi target populasi dari survei ini adalah pelaku usaha yang terdaftar dalam administrasi Direktorat Jenderal Pajak (DJP) sebagai Wajib Pajak

³ Sebagai sebuah *predictive model*, ANN dapat dikombinasikan dengan model *Decision Tree* untuk analisis lebih lengkap (IBM 2019). Salah satu pendekatan yang populer dalam model *Decision Tree* adalah *Chi-Squared Automatic Interaction Detection* (CHAID). Dalam pendekatan CHAID, hanya ada tiga tingkatan *tree*—yang dalam hal ini merujuk pada variabel independen—yang dihasilkan (IBM 2017). Atas dasar pertimbangan ini, penulis hanya fokus pada tiga variabel terpenting.

⁴ Survei dilakukan selama tiga minggu (21 Juli sampai dengan 7 Agustus 2020). Survei ini merupakan bagian kegiatan dari Kelompok Kerja Penerimaan Program Pemulihan Ekonomi Nasional (Pokja Penerimaan PEN) Kementerian Keuangan Republik Indonesia yang diketuai oleh Staf Ahli Menteri Keuangan Bidang Kepatuhan Pajak.

strategis.⁵ Survei dilakukan secara daring dengan cara mengirimkan surat elektronik (surel) atau *e-mail* ke alamat surel Wajib Pajak strategis, dengan mencantumkan tautan survei. Survei bersifat anonim; responden tidak diberikan pertanyaan apapun yang berkaitan dengan identitas individu atau entitas.⁶

Berdasarkan ketersediaan data hasil survei, terdapat lima karakteristik usaha yang secara khusus dianalisis dalam penelitian ini: (i) *jumlah pekerja*; (ii) *pangsa pasar utama*; (iii) *omzet setahun*; (iv) *sifat usaha utama*; dan (v) *sumber utama pasokan*. Variabel *jumlah pekerja* diukur dengan skala ordinal dan memiliki tujuh pilihan: (i) satu orang/pemilik, (ii) 2 – 10 orang, (iii) 11 – 50 orang; (iv) 51 – 100 orang, (v) 101 – 250 orang, (vi) 251 – 500 orang, dan (vii) di atas 500 orang. Variabel *pangsa pasar utama* merupakan variabel kategorikal dan memiliki tiga pilihan: (i) lokal, (ii) ekspor, dan (iii) campuran lokal dan ekspor.

Selanjutnya, variabel *omzet tahunan* diukur dengan skala ordinal dalam enam pilihan: (i) kurang dari Rp 5 miliar, (ii) antara Rp 5 miliar – Rp 10 miliar, (iii) antara Rp 10 miliar – Rp 25 miliar; (iv) antara Rp 25 miliar – Rp 50 miliar, (v) antara Rp 50 miliar – Rp 100 miliar, dan (vi) di atas Rp 100 miliar. Variabel *sifat usaha utama* merupakan variabel kategorikal yang memiliki dua pilihan: (i) produsen dan (ii) non-produsen. Terakhir,

⁵ Wajib Pajak strategis adalah Wajib Pajak orang pribadi dan Wajib Pajak badan berstatus pusat yang terdaftar di Kantor Pelayanan Pajak (KPP) di lingkungan Kantor Wilayah (Kanwil) DJP Wajib Pajak Besar, KPP di lingkungan Kanwil DJP Jakarta Khusus, seluruh KPP Madya, dan Wajib Pajak dengan kriteria tertentu—umumnya terkait dengan besarnya nilai pembayaran pajak—yang terdaftar di tiap-tiap KPP Pratama. Secara total terdapat 352 KPP di seluruh Indonesia. Secara umum terdapat sekitar 300 s.d. 500 Wajib Pajak strategis dalam setiap KPP. Dengan demikian, pengertian pelaku usaha dalam penelitian ini hanya merujuk pelaku usaha yang teradministrasi dalam sistem perpajakan DJP.

⁶ Hasil uji keterwakilan (*representativeness tests*) responden survei menunjukkan bahwa responden survei adalah representasi yang sangat baik dari target populasi baik dari segi sebaran lokasi usaha, sebaran sektor usaha, dan sebaran omzet tahunan.

variabel *sumber utama pasokan* merupakan variabel kategorikal dengan empat pilihan skala nominal: (i) dalam negeri, (ii) luar negeri (impor), (iii) campuran dalam dan luar negeri, dan (iv) jasa murni (tidak membutuhkan bahan baku/barang dagangan).

2.2 Metode Analisis

Merujuk pada tujuan penelitian menurut Babbie (2010), fokus tulisan ini adalah memberikan gambaran (*description*), bukan penjelasan (*explanation*). Untuk mencapai tujuan ini, studi ini mengadopsi pendekatan *applied research*. *Applied research* didesain untuk menjawab permasalahan tertentu atau menawarkan solusi dari permasalahan praktis yang terjadi (Neuman 2011).⁷ Seperti yang telah dijelaskan sebelumnya, keikutsertaan pelaku usaha dalam insentif pajak, secara intuisi, banyak berkaitan dengan faktor-faktor yang bersifat non-linier. Dalam area yang bersifat non-linier, tingkat akurasi dari metode yang ada dalam *artificial intelligence* cenderung lebih unggul dari pada metode statistik tradisional (Bahrammirzaee 2010).⁸ Berdasarkan pertimbangan ini—dan merujuk pada tujuan studi, penulis menggunakan pendekatan *Artificial Neural Network* (ANN).⁹

⁷ Implikasinya, *applied social research* jarang memiliki keterkaitan kuat dengan aktivitas membangun, menguji, atau menghubungkan teori secara mendalam (Neuman 2011).

⁸ Model regresi linier memiliki struktur model yang kaku dan serangkaian asumsi yang diterapkan sebelum belajar dari data. Sebaliknya, *neural network* dapat memperkirakan berbagai model statistik tanpa mengharuskan kita berhipotesis terlebih dahulu mengenai hubungan tertentu antara variabel dependen dan independen—karena bentuk hubungan ditentukan selama model melakukan proses *learning*. *Neural network* akan secara fleksibel memilih apakah model linier atau non-linier dalam proses ini. Konsekuensinya, sebagai akibat fleksibilitas ini, *synaptic weight* dari *neural network* tidak mudah diinterpretasikan. Jika penafsiran dari hubungan linier antara variabel dependen dan independen merupakan tujuan utama, maka pendekatan model statistik yang tradisional lebih disarankan (IBM 2019).

⁹ ANN adalah salah satu metode *artificial intelligence* yang paling populer digunakan di bidang keuangan (Bahrammirzaee 2010).

ANN tersusun dari metode matematika yang banyak digunakan dalam penelitian prediksi dan klasifikasi (Aryadoust & Baghaei 2016). ANN adalah pendekatan yang banyak digunakan untuk analisis *predictive data mining* karena memiliki tingkat akurasi, fleksibilitas, dan kemudahan penggunaan—khususnya dalam situasi di mana proses yang mendasarinya rumit (IBM 2019). Keunggulan utama dari ANN adalah pengenalan dan sekaligus klasifikasi pola (*pattern*) karena sifat pembelajaran adaptif non-parametrik non-linier yang dimiliki (Zhang et al. 1999). ANN dibangun pada premis bahwa hubungan yang tepat antara variabel independen dan variabel dependen dapat diestimasi menggunakan fungsi matematika non-linier (Aryadoust & Baghaei 2016).¹⁰

Lebih spesifik, dalam studi ini penulis menggunakan modul *Multilayer Perceptron* (MLP) dari IBM SPSS.¹¹ Konsepsi MLP dalam ANN beranjak dari *backpropagation learning error*—sebuah algoritma yang paling sering digunakan dalam ANN (Pérez López et al. 2019). Pendekatan MLP lebih populer digunakan dalam ANN dibanding pendekatan lain (Zhang et al. 1999). Studi yang dilakukan oleh Sánchez-Serrano et al. (2020) juga menemukan bahwa metode MLP menghasilkan tingkat akurasi yang lebih tinggi daripada pendekatan *Radial Basis Function* (RBF).¹²

¹⁰ Pembahasan mengenai ANN melibatkan model matematis yang cukup kompleks. Pembaca dapat merujuk pada Ripley (1996) dan Haykin (1999) untuk pembahasan model teoritis dan matematis yang komprehensif. Pembaca yang tertarik aspek aplikatif dari ANN ini dapat mempelajari lebih lanjut dalam ‘*IBM SPSS Neural Networks 26*’ (IBM 2019).

¹¹ Terdapat dua modul aplikasi prediksi berbasis ANN dalam IBM SPSS: MLP dan *Radial Basis Function* (RBF) (IBM 2019). MLP dan RBF adalah dua arsitektur *neural network* yang paling banyak digunakan (Ripley 1996)

¹² Perlu dicatat, metode MLP tidak selalu lebih akurat. Misalnya, Jupri dan Sarno (2018) melakukan komparasi empat algoritma klasifikasi C4.5, *Support Vector Machine* (SVM), *K-Nearest Neighbour* (KNN) dan MLP

Merujuk pada formasi partisi yang dilakukan oleh Bekesiene et al. (2021), data yang dianalisis diklasifikasikan secara acak ke dalam tiga kelompok: (i) data *training* sebanyak 60%, (ii) data *testing* sebanyak 20%, dan (iii) data *holdout* sebanyak 20%—atau jamak disebut formasi 60%-20%-20%.¹³ Data *training* digunakan untuk mencari bobot dan membangun model. Data *testing* digunakan untuk menemukan kesalahan dan mencegah *overtraining* selama mode pelatihan. Sedangkan data *holdout* digunakan untuk memvalidasi model (IBM 2019).¹⁴

3 HASIL DAN PEMBAHASAN

3.1 Hasil Analisis

3.1.1 Statistik deskriptif

Terdapat 12.361 data pelaku usaha dari seluruh Indonesia yang dianalisis dalam studi ini. Komposisi jumlah pelaku usaha berdasarkan lokasi di enam pulau utama adalah sebagai berikut: Jawa (70%, n=8.681), Sumatra (13%; n=1.633), Kalimantan (6%; n=706), Sulawesi (5%; n=612), Bali dan Nusa Tenggara (5%; n=573), Papua dan Maluku (1%; n=156). Statistik deskriptif mengenai variabel yang diuji dapat dilihat di panel A Tabel 1. Panel B Tabel 1 menyajikan koefisien korelasi (r) dari variabel yang diuji. Nilai koefisien korelasi paling besar dihasilkan oleh hubungan antara variabel *jumlah pekerja* dan *omzet setahun* ($r = 0,644$).

untuk mengklasifikasi tingkat kepatuhan dari Wajib Pajak dan menyimpulkan bahwa algoritma C4.5 merupakan algoritma klasifikasi yang paling akurat.

¹³ Studi yang dilakukan oleh Bekesiene et al. (2021) menunjukkan bahwa format partisi 60%-20%-20% lebih optimal dibandingkan dengan format 50%-30%-20% maupun 70%-20%-10%.

¹⁴ Formasi partisi yang cukup jamak dilakukan adalah 70% untuk *data training* dan 30% untuk *data testing* (70%-30%). Formasi ini biasanya digunakan untuk observasi yang relatif tidak banyak (misalnya kurang dari 1.000) dan oleh karenanya tidak memiliki alokasi data yang digunakan untuk memvalidasi model, yaitu data *holdout*.

Tabel 1: Descriptive Statistics dan Hubungan antar Variabel yang Diuji

Panel A. Descriptive statistics

	N	Min.	Max.	Mean	SD	Variance
Status pemanfaatan insentif	12.361	0	1	0,49	0,500	0,250
Jumlah pekerja	12.361	1	7	3,11	1,475	2,176
Sifat usaha utama	12.361	1	2	1,26	0,439	0,193
Sumber utama pasokan	12.361	1	4	2,04	1,212	1,468
Pangsa pasar utama	12.361	1	3	1,25	0,634	0,402
Omzet setahun	12.361	1	6	2,59	1,744	3,042
Valid N (listwise)	12.361					

Panel B. Hubungan antar Variabel yang Diuji

	Status pemanfaatan insentif	Jumlah pekerja	Sifat usaha utama	Sumber utama pasokan	Pangsa pasar utama	Omzet setahun
N	12.361	12.361	12.361	12.361	12.361	12.361
Status pemanf. insentif	1	.356**	.240**	.019*	.179**	.345**
Jumlah pekerja	.356**	1	.382**	.072**	.288**	.644**
Sifat usaha utama	.240**	.382**	1	-.088**	.233**	.266**
Sumber utama pasokan	.019*	.072**	-.088**	1	.143**	.024**
Pangsa pasar utama	.179**	.288**	.233**	.143**	1	.219**
Omzet setahun	.345**	.644**	.266**	.024**	.219**	1

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Meski demikian, kekuatan hubungan antara dua variabel ini bersifat sedang (*moderate*) karena memiliki nilai $r < 0,69$ (Schober et al. 2018).¹⁵ Hubungan korelasional status pemanfaatan insentif dengan variabel *jumlah pekerja*, *sifat usaha utama*, *sumber pasokan utama*, dan *pangsa pasar utama* memiliki nilai koefisien yang bervariasi dengan kecenderungan tingkat hubungan yang lemah karena memiliki nilai $< 0,40$.

¹⁵ Menurut Schober et al. (2018), koefisien korelasi 0,40-0,69 menunjukkan korelasi sedang, 0,70-0,89 menunjukkan korelasi kuat, dan 0,90-1,00 menunjukkan korelasi yang sangat kuat.

3.1.2 Artificial Neural Network (ANN)

Tujuan pertama dari studi ini adalah untuk menguji seberapa akurat pendekatan *multilayer perceptron* (MLP) *Neural Network* dapat memprediksi keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak (memanfaatkan atau tidak memanfaatkan) dengan menganalisis data lima karakteristik utama dari pelaku usaha dan mengetahui urutan tingkat kepentingan dari variabel bebasnya.

Tabel 2 menunjukkan bahwa sebanyak 12.361 observasi digunakan untuk membangun model ANN dalam studi ini. Seperti yang telah dijelaskan sebelumnya, mengikuti hasil studi yang dilakukan oleh Bekesiene et al. (2021), data pelaku usaha hasil survei dalam studi ini dibagi dalam tiga kelompok dengan proporsi: (i) data *training* sebanyak 7.415 observasi (60%); (ii) data *testing* sebanyak 2.460 observasi (19,9%); dan (iii) data *holdout* sebanyak 2.486 observasi (20,1%). Tidak ada observasi yang dikeluarkan (*excluded*) oleh aplikasi dalam analisis ini.

Tabel 2: Case Processing Summary

		N	Percent
Sample	Training	7415	60.0%
	Testing	2460	19.9%
	Holdout	2486	20.1%
Valid		12361	100.0%
Excluded		0	
Total		12361	

Tabel 3 menunjukkan jumlah *neuron* dalam tiap *layer* dan lima variabel bebas yang digunakan dalam analisis (*input layer*): (i) *sifat usaha utama*; (ii) *pangsa pasar utama*; (iii) *sumber utama pasokan*; (iv) *jumlah pekerja*; dan (v) *omzet setahun*. Dalam analisis

ini tiga variabel kategorikal masuk dalam kelompok *factors* dan dua variabel ordinal masuk dalam kategori *covariates*, dengan jumlah total sebelas unit.

Tabel 3: Network Information

Input Layer	Factors	1	Sifat usaha utama
		2	Pangsa pasar utama
		3	Sumber utama pasokan
	Covariates	1	Jumlah pekerja
		2	Omset setahun
	Number of Units ^a		
Rescaling Method for Covariates		Standardized	
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1 ^a		8
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Apakah sudah/sedang memanfaatkan insentif?
	Number of Units		2
	Activation Function		Softmax
	Error Function		Cross-entropy

a. Excluding the bias unit

Fitur *automatic architecture* dari aplikasi menunjukkan ada delapan unit di *hidden layers* yang terbentuk, sementara untuk *output layer* terdapat dua unit yang merepresentasikan status pemanfaatan insentif pajak.¹⁶ Aktivasi fungsi *hidden layer* dalam analisis ini menggunakan *hyperbolic tangent*, sementara untuk *output layer* menggunakan *Softmax*. *Cross-entropy* digunakan sebagai *error function* karena penggunaan metode *Softmax* sebagai fungsi aktivasi.

Grafik 1 menunjukkan diagram jaringan yang digunakan SPSS untuk memprediksi keikutsertaan pelaku usaha (*tidak ikut insentif=0, ikut insentif=1*) berdasarkan lima

¹⁶ *Hidden layers* memungkinkan ANN untuk melakukan emulasi terhadap pola non-linier dalam data secara lebih akurat. Tanpa *hidden layer*, ANN akan berperilaku seperti model linier biasa yang tidak memiliki kemampuan untuk mendeteksi pola non-linier (Aryadoust & Baghaei 2016).

karakteristik utama dari pelaku usaha: (i) jumlah pekerja; (ii) pangsa pasar utama; (iii) omzet setahun; (iv) sifat utama usaha; dan (v) sumber utama pasokan.

Grafik 1: Network Diagram

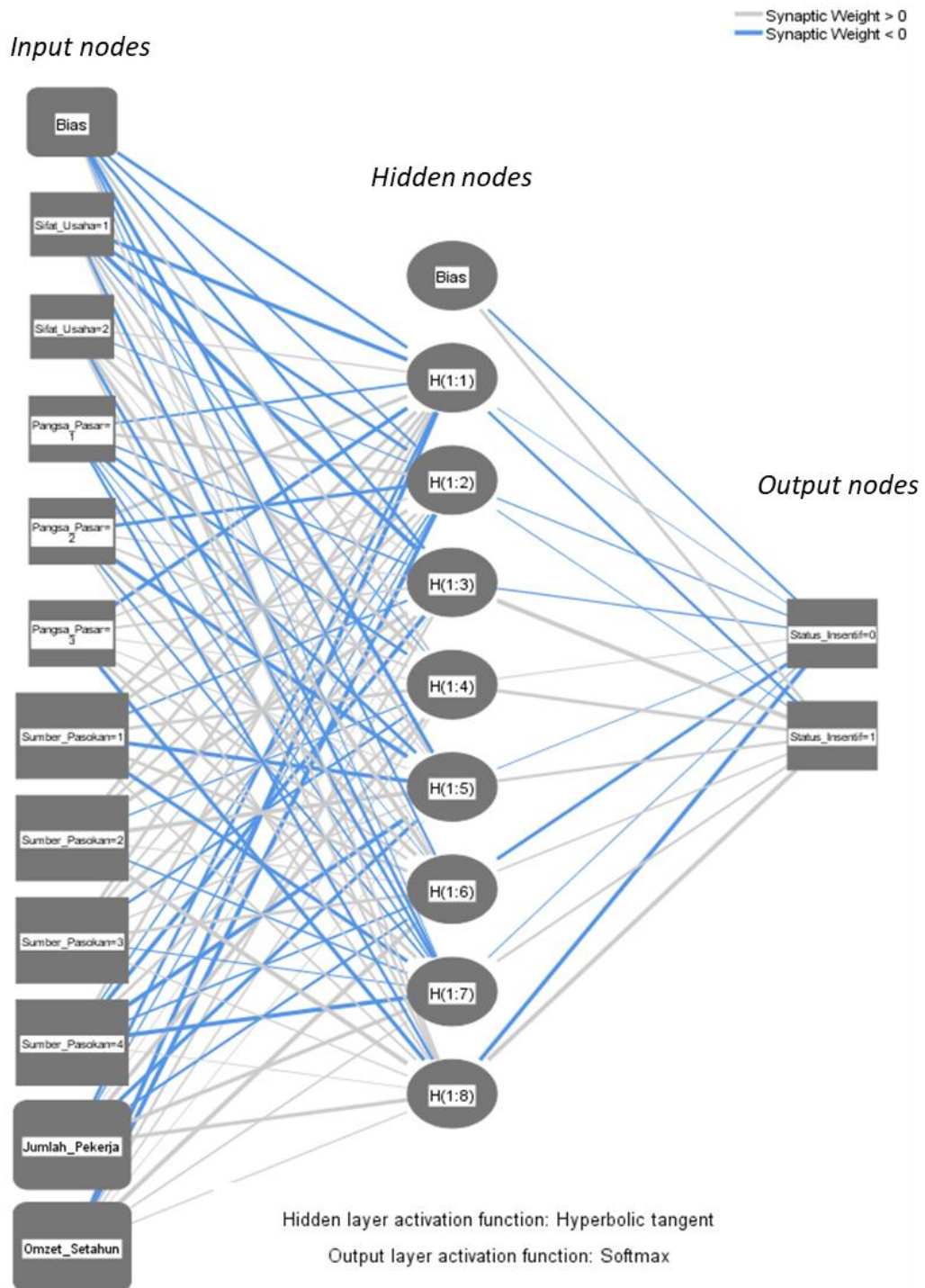


Diagram jaringan di Grafik 1 menunjukkan terdapat sebelas *input nodes*, delapan *hidden nodes*, dan dua *output nodes* yang menunjukkan status keikutsertaan pelaku usaha dalam insentif pajak.¹⁷

Selanjutnya, Tabel 4 menyajikan ringkasan informasi yang berkaitan dengan hasil *training* (dan *testing*) dan hasil uji terhadap *holdout sample*. Nilai dari *cross-entropy error* juga disajikan baik untuk *training sample* maupun *testing sample* karena ini menunjukkan nilai *error function* yang diminimalkan oleh model ANN selama fase *training*. Nilai *cross-entropy error* yang lebih kecil (1.449) untuk *testing sample* dibandingkan *training sample* (4.472) mengindikasikan bahwa tidak terjadi *overfitted* terhadap data *training* dalam model jaringan.

Tabel 4: Model Summary

Training	Cross Entropy Error	4472.528
	Percent Incorrect Predictions	32.6%
	Stopping Rule Used	10 consecutive step(s) with no decrease in error ^a
	Training Time	0:00:00,35
Testing	Cross Entropy Error	1449.531
	Percent Incorrect Predictions	30.3%
Holdout	Percent Incorrect Predictions	31.1%

Dependent Variable: Apakah sudah/sedang memanfaatkan insentif?

a. Error computations are based on the testing sample.

Hasil ini memberikan justifikasi mengenai peran dari *testing sample* yaitu mencegah terjadinya *overtraining*. Berdasarkan informasi dalam Tabel 4, proporsi dari prediksi

¹⁷ Struktur ini dikenal juga sebagai '*feedforward architecture*' karena hubungan dalam jaringan bergerak maju dari lapisan *input* menuju lapisan *output* tanpa *feedback loop*. Lapisan *input* terdiri dari prediktor, lapisan *hidden* terdiri dari unit yang tidak dapat terobservasi (*unobservable*), dan lapisan *output* berisi respon. Unit dalam *output* merupakan kumpulan dari beberapa fungsi dari unit di lapisan *hidden* (IBM 2019).

yang tidak tepat dari *training sample* dan *testing sample* masing-masing adalah 32,6% dan 30,3%. Untuk *holdout sample*, proporsi prediksi yang tidak tepat adalah 31,1%.

Tabel 5 menyajikan besaran dari *synaptic weights* antara tiga *layer* seperti terlihat dalam Grafik 1. *Synaptic weights* adalah estimasi koefisien yang menunjukkan hubungan antara unit pada suatu *layer* dengan unit pada *layer* berikutnya. Estimasi dari *synaptic weights* ini hanya didasarkan pada data *training*,¹⁸ dan oleh karenanya umumnya tidak digunakan untuk menginterpretasikan hasil uji ANN (IBM 2019).

Tabel 5: Parameter Estimates

Predictor	Predicted									Output Layer	
	Hidden Layer 1									[Status_Insentif=0]	[Status_Insentif=1]
	H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	H(1:8)			
Input Layer	(Bias)	-.300	-.211	-.266	.240	-.363	-.207	-.039	.168		
	[Sifat_Usaha=1]	-.427	-.316	-.535	-.054	.270	.057	-.284	.212		
	[Sifat_Usaha=2]	.116	-.052	.047	.399	.023	.492	-.275	.246		
	[Pangsa_Pasar=1]	-.212	.342	-.161	-.257	-.305	-.012	-.144	-.250		
	[Pangsa_Pasar=2]	.353	-.350	.118	.204	-.442	.398	-.170	.200		
	[Pangsa_Pasar=3]	-.405	.139	.189	.036	.042	.230	.268	-.372		
	[Sumber_Pasokan=1]	.352	.346	-.179	.308	-.413	.315	-.360	.167		
	[Sumber_Pasokan=2]	.317	.255	.418	-.105	.532	.040	-.152	.767		
	[Sumber_Pasokan=3]	.324	.166	.374	-.235	.143	.300	-.055	.088		
	[Sumber_Pasokan=4]	.255	-.131	-.410	.266	-.479	-.191	-.529	.019		
	Jumlah_Pekerja	-.296	-.079	.012	.449	-.366	-.231	.402	.394		
Omzet_Setahun	-.477	-.545	.225	.119	.008	.565	.151	.071			
Hidden Layer 1	(Bias)									-.195	.265
	H(1:1)									-.005	-.247
	H(1:2)									-.063	-.029
	H(1:3)									-.093	.846
	H(1:4)									.042	.347
	H(1:5)									-.022	.249
	H(1:6)									-.342	.158
	H(1:7)									-.030	.219
	H(1:8)									-.407	.530

Tabel 6 menyajikan informasi mengenai tingkat akurasi dari model ANN untuk sampel data *training*, *testing*, dan *holdout*. Masing-masing kelompok data memiliki tingkat

¹⁸ Meskipun data dipartisi dalam tiga kategori, *training*, *testing* dan *holdout*, estimasi *synaptic weights* hanya didasarkan pada data *training* (IBM 2019).

akurasi secara berurutan 67,4%, 69,7%, dan 68,9%. Secara agregat, pelaku usaha yang diprediksi tidak ikut insentif adalah sebesar 56,3% untuk data *training*, 58,8% untuk data *testing*, dan 57,2% untuk data *holdout*. Sedangkan pelaku usaha yang ikut insentif adalah 43,7% untuk data *training*, 41,2% untuk data *testing*, dan 42,8% untuk data *holdout*. Tingkat akurasi klasifikasi dari masing-masing kelompok data mendekati 70%: data *training* (67,4%), data *testing* (69,7%), dan data *holdout* (68,9%).

Lebih spesifik, dalam data *training* 2.758 pelaku usaha yang tidak ikut insentif terklasifikasi secara akurat (*true negative*) dan terdapat 999 pelaku usaha yang tidak ikut insentif namun diprediksi sebagai ikut insentif (*false positive*), menghasilkan tingkat *sensitivity* 73,4%. Dalam kelompok data ini terdapat 2.238 pelaku usaha yang terklasifikasi secara akurat ikut insentif (*true positive*), namun terdapat 1.420 pelaku usaha ikut insentif namun terklasifikasi sebagai tidak ikut insentif (*false negative*). Hal ini menghasilkan tingkat *specificity* 61,2%.

Tabel 6: Accuracy of Classification

Sample	Observed	Predicted		Percent Correct
		Tidak	Ya	
Training	Tidak	2758	999	73.4%
	Ya	1420	2238	61.2%
	Overall Percent	56.3%	43.7%	67.4%
Testing	Tidak	967	266	78.4%
	Ya	480	747	60.9%
	Overall Percent	58.8%	41.2%	69.7%
Holdout	Tidak	957	307	75.7%
	Ya	465	757	61.9%
	Overall Percent	57.2%	42.8%	68.9%

Dependent Variable: Apakah sudah/sedang memanfaatkan insentif?

Selanjutnya, dalam data *testing* 967 pelaku usaha yang tidak ikut insentif terklasifikasi secara akurat (*true negative*) dan terdapat 226 pelaku usaha yang tidak ikut insentif

namun diprediksi sebagai ikut insentif (*false positive*), menghasilkan tingkat *sensitivity* 69,7%. Dalam kelompok data ini terdapat 747 pelaku usaha yang terklasifikasi secara akurat ikut insentif (*true positive*), namun terdapat 480 pelaku usaha yang ikut insentif namun terklasifikasi sebagai tidak ikut insentif (*false negative*). Komposisi ini menghasilkan tingkat *specificity* 60,9%.

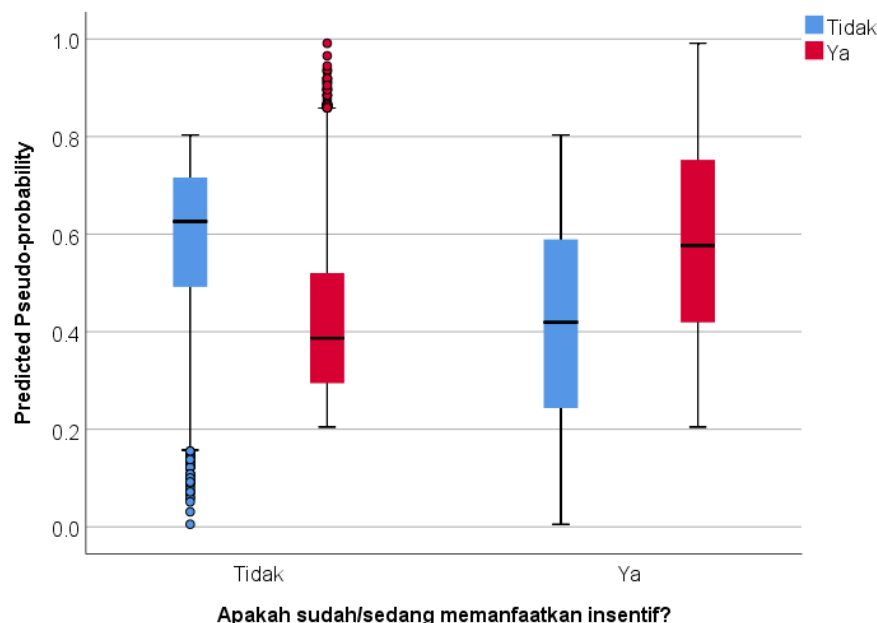
Terakhir, dalam data *holdout* 957 pelaku usaha yang tidak ikut insentif terklasifikasi secara akurat (*true negative*) dan terdapat 266 pelaku usaha yang tidak ikut insentif namun diprediksi sebagai ikut insentif (*false positive*), menghasilkan tingkat *sensitivity* 75,7%. Dalam kelompok data ini terdapat 757 pelaku usaha yang terklasifikasi secara akurat ikut insentif (*true positive*), namun terdapat 465 pelaku usaha yang ikut insentif namun terklasifikasi sebagai tidak ikut insentif (*false negative*). Hal ini menghasilkan tingkat *specificity* 61,9%.

Aplikasi IBM SPSS juga menampilkan grafik *predicted pseudo-probability* untuk dua kelompok pelaku usaha dari status pemanfaatan insentif dalam sebuah diagram *box-plot*. Grafik ini secara spesifik menggambarkan prediksi dari dua kategori dari variabel tergantung status pemanfaatan insentif. Perlu dicatat, grafik ini menunjukkan dua *box-plot* yang mewakili *predicted pseudo-probability* berdasarkan hasil analisis dari seluruh data ($n=12.361$). Untuk setiap *box-plot* di masing-masing kategori, nilai di atas 0,5 mengindikasikan prediksi yang akurat. Sebaliknya, nilai di bawah 0,5 menunjukkan prediksi yang tidak akurat.

Penafsiran yang lebih detail dari Grafik 2 dapat dimulai dengan melihat bagian kiri dari grafik. Sisi kiri dari sumbu-x menunjukkan kategori pelaku usaha yang tidak

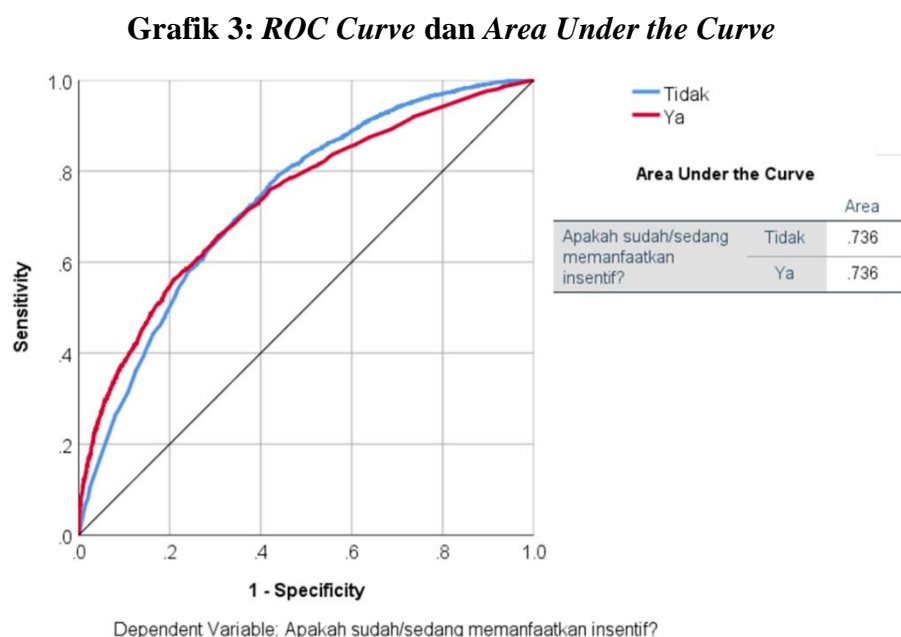
memanfaatkan insentif. Terdapat dua *box-plot* di kategori ini: pemanfaat insentif dan non-pemanfaat. Selanjutnya, sumbu-y menunjukkan nilai *predicted pseudo-probability*. Grafik 2 menggambarkan prediksi yang cukup akurat untuk pelaku usaha yang tidak ikut insentif karena mayoritas area *box-plot* dari kelompok ini berada di atas angka 0,5 untuk kategori non-pemanfaat dan berada di bawah angka 0,5 untuk kategori pemanfaat insentif. Meski demikian, seperti terlihat di sisi kanan dari grafik, tingkat akurasi prediksi untuk pelaku usaha pemanfaat insentif lebih rendah dibandingkan dengan tingkat akurasi prediksi dari pelaku usaha non-pemanfaat. Hal ini dapat terlihat dari luas area yang berada di atas nilai 0,5 untuk kategori pelaku usaha non-pemanfaat insentif yang menunjukkan keberadaan *false positive*. Di sisi lain, luas area yang berada di bawah nilai 0,5 untuk kategori pelaku usaha pemanfaat insentif juga cukup jelas terlihat. Hal ini menunjukkan keberadaan *false negative*.

Grafik 2: Predicted-by-observed Chart



Grafik 3 menunjukkan kurva *Receiver Operating Characteristic* (ROC). Kurva ROC adalah penggambaran kinerja klasifikasi dua dimensi (Fawcett 2006). Grafik ini memberikan gambaran tingkat *sensitivity* dan *specificity* berdasarkan gabungan antara data *training* dan data *sampling*. Garis diagonal 45-derajat dari bagian kiri bawah ke bagian atas kanan menunjukkan garis *no-discrimination*. Titik yang berada di bawah di bawah garis *no-discrimination* mengindikasikan klasifikasi yang tidak akurat, dan titik di atas garis *no-discrimination* menunjukkan hasil klasifikasi yang efektif (Fawcett 2006).

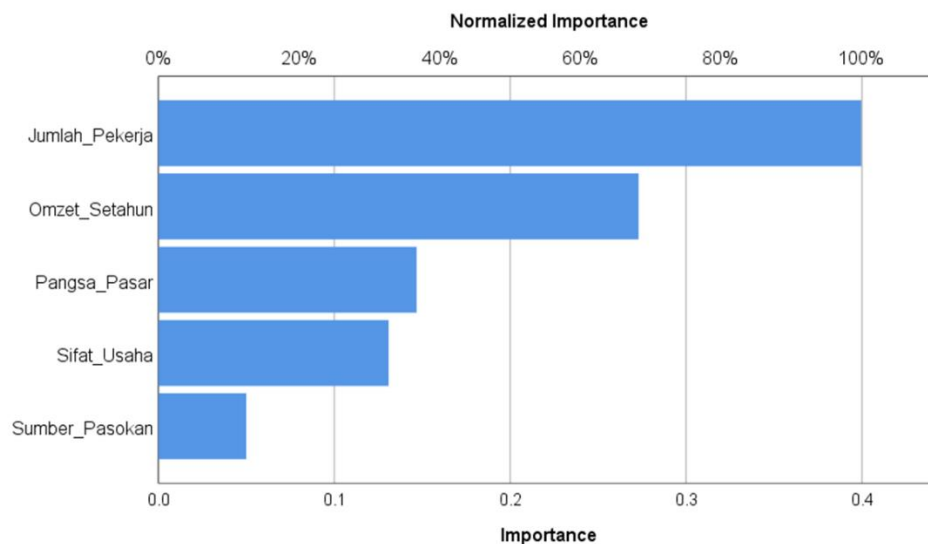
Dalam Grafik 3 juga disajikan nilai *area under the curve* (AUC) sebesar 0,74. Nilai AUC sebesar 0,74 mengindikasikan bahwa jika satu pelaku usaha dari kelompok pemanfaat insentif dan satu pelaku usaha dari kelompok non-pemanfaat insentif terpilih secara acak, maka terdapat 74% kemungkinan bahwa model *predicted pseudo-probability* untuk pelaku usaha yang pertama masuk dalam kategori pemanfaat insentif (status insentif = 1).



Secara umum, nilai AUC diklasifikasikan dalam lima kategori: 0,50–0,60 (gagal); 0,60–0,70 (kurang akurat); 0,70–0,80 (cukup akurat); 0,80–0,90 (akurat); dan 0,90–1,00 (sangat akurat). Dengan demikian, nilai AUC dari studi menunjukkan bahwa hasil prediksi ini cukup akurat.

Terkait dengan tujuan kedua dari studi ini, Grafik 4 menunjukkan seberapa besar pengaruh variabel independen dalam model ANN, yang diukur berdasarkan tingkat kepentingan relatif dan tingkat kepentingan yang dinormalisasi (*normalised importance*).

Grafik 4: Independent Variable Importance



Independent Variable Importance		
	Importance	Normalized Importance
Jumlah pekerja	.399	100.0%
Pangsa pasar utama	.147	36.7%
Omset setahun	.273	68.3%
Sifat usaha utama	.131	32.8%
Sumber utama pasokan	.050	12.5%

Seperti terlihat dalam grafik, variabel *jumlah pekerja* memiliki skor tertinggi (0,399; *normalised importance* = 100%), disusul oleh *omzet setahun* (0,273; *normalised*

importance = 68,3%) dan *pangsa pasar* (0,147; *normalised importance* = 36,7%). Dua variabel dengan pengaruh terendah adalah *sifat usaha* (0,131; *normalised importance* = 32,8%) dan *sumber pasokan* (0,050; *normalised importance* = 12,5%).

Dari Grafik 4 dapat disimpulkan bahwa jumlah pekerja yang dimiliki pelaku usaha, besaran omzet tahunan, dan pangsa pasar utama dari pelaku usaha memiliki efek yang paling besar dalam bagaimana ANN mengklasifikasi dan memprediksi keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak.

3.1.3 *Robustness Tests*

Untuk menguji seberapa *robust* temuan sebelumnya, penulis melakukan *sampling variability* dan mereplikasi pendekatan ANN dengan menguji data berdasarkan lokasi pelaku usaha di enam kelompok pulau. Uji *robustness* ini juga berkaitan dengan tujuan ketiga dari studi ini yaitu untuk mengetahui seberapa kuat (*robust*) hasil prediksi ANN jika sampel data yang digunakan berbasis lokasi pelaku usaha. Ringkasan hasil analisis akhir berupa nilai *variable independent importance* dari masing-masing kelompok pulau kemudian dikomparasi dan disajikan di Tabel 7.¹⁹

Tabel 7 menunjukkan bahwa variabel *jumlah pekerja* konsisten memiliki skor *Independent Variable Importance* (IVI) yang paling besar di seluruh kelompok pulau: Jawa (0,383; *normalised importance* = 100%), Sumatra (0,402; *normalised importance* = 100%), Kalimantan (0,465; *normalised importance* = 100%), Sulawesi (0,275; *normalised*

¹⁹ Dengan tujuan menjaga keringkasan tulisan, penjelasan teknis dan detail terkait hasil analisis seperti yang dijelaskan sebelumnya tidak dibahas dalam bagian *robustness tests* ini.

importance = 100%), Bali – Nusa Tenggara (0,493; *normalised importance* = 100%), dan Papua – Maluku (0,410; *normalised importance* = 100%).²⁰

Tabel 7: Komparasi Tiga Terbesar Nilai *Independent Variable Importance* Berdasarkan Lokasi Pelaku Usaha

	Jawa (n=8.681)	Sumatra (n=1.633)	Kalimantan (n=706)	Sulawesi (n=612)	Bali – NT (n=573)	Papua – Maluku (n=156)
Jumlah pekerja	0,383 (100%)	0,402 (100%)	0,465 (100%)	0,275 (100%)	0,493 (100%)	0,410 (100%)
Omzet setahun	0,333 (86,9%)	0,268 (66,5%)	0,119 (25,7%)	0,255 (92,8%)	0,149 (30,2%)	0,291 (71%)
Pangsa pasar utama	0,090 (23,6%)	0,124 (30,7%)	0,240 (51,6%)	0,228 (83,1%)	0,225 (45,6%)	0,089 (21,8%)
Sifat usaha utama	0,170 (44,3%)	0,117 (29%)	0,104 (22,4%)	0,169 (61,4%)	0,052 (10,5%)	0,122 (29,7%)
Sumber utama pasokan	0,024 (6,2%)	0,090 (22,3%)	0,072 (15,4%)	0,073 (26,6%)	0,082 (16,6%)	0,088 (21,5%)

Catatan: Persentase dalam kurung menunjukkan nilai yang dinormalisasi (*normalised importance*). Angka dalam kotak menunjukkan deviasi posisi menurut hasil prediksi ANN menggunakan sampel data nasional (n=12.361). Variabel disusun secara berurutan berdasarkan urutan di Grafik 4.

Variabel *omzet setahun* menduduki posisi kedua dalam skor IVI di empat kelompok pulau, kecuali pulau Kalimantan, dan Bali – Nusa Tenggara. Nilai IVI dari variabel *omzet setahun* untuk kelompok pulau Jawa, Sumatra, Sulawesi, dan Papua – Maluku masing-masing secara berurutan sebesar 0,333 (*normalised importance* = 86,9%), 0,268 (*normalised importance* = 66,5%), 0,255 (*normalised importance* = 92,8%), dan 0,291 (*normalised importance* = 71%). Sedangkan untuk kelompok pulau Kalimantan dan Bali

²⁰ Nilai *normalized* IVI dalam bentuk persentase disajikan untuk menunjukkan peringkat dari variabel ini secara relatif terhadap variabel lain yang dikomparasi. Nilai 100% menunjukkan bahwa nilai IVI dari variabel ini adalah paling besar dibandingkan variabel yang lain.

– Nusa Tenggara, masing-masing nilai IVI dari variabel *omzet setahun* adalah 0,119 (*normalised importance* = 25,7%) dan 0,149 (*normalised importance* = 30,2%).

Terakhir, variabel *pangsa pasar utama* menduduki posisi ketiga dalam skor IVI di dua kelompok pulau—Sumatra (0,124; *normalised importance* = 30,7%) dan Sulawesi (0,228; *normalised importance* = 83,1%) dan posisi kedua di Kalimantan (0,240; *normalised importance* = 51,6%) dan Bali – Nusa Tenggara (0,225; *normalised importance* = 45,6%). Yang menarik, untuk kelompok pulau Jawa dan Papua – Maluku, variabel *pangsa pasar utama* menduduki posisi keempat dengan besaran IVI secara berurutan 0,090 (*normalised importance* = 23,6%) dan 0,089 (*normalised importance* = 21,8%). Di dua kelompok pulau ini variabel *sifat usaha utama* menduduki posisi ketiga dengan besaran IVI: pulau Jawa (0,170; *normalised importance* = 44,3%) dan pulau Bali – Nusa Tenggara (0,122; *normalised importance* = 29,7%). Hasil analisis *robustness test* menunjukkan bahwa faktor penentu keikutsertaan pemanfaatan insentif pajak oleh pelaku usaha homogen di seluruh kelompok pulau, khususnya untuk variabel *jumlah pekerja*.

3.2 Pembahasan

Hasil analisis menunjukkan bahwa tingkat akurasi dari pendekatan ANN dalam studi ini sekitar 70%. Hal ini sesuai dengan kesimpulan Bahrammirzaee (2010) yang menyatakan bahwa studi empiris komparatif menunjukkan bahwa keberhasilan penggunaan ANN di bidang keuangan sangat menjanjikan, meskipun tidak bisa dibilang sangat akurat. Misalnya, hasil uji empiris yang dilakukan Lin et al. (2012) menunjukkan bahwa pendekatan *decision tree* memiliki akurasi yang lebih tinggi

untuk mendeteksi perilaku ketidakpatuhan pajak dari Wajib Pajak orang pribadi, sementara ANN mendeteksi secara lebih akurat untuk Wajib Pajak badan.

Hasil studi ini juga menunjukkan adanya kecenderungan bahwa tingkat akurasi prediksi untuk pelaku usaha yang tidak ikut insentif (*true negative*) lebih tinggi dibandingkan dengan tingkat akurasi untuk pelaku usaha yang ikut insentif (*true positive*). Secara berurutan, tingkat akurasi dari prediksi *true negative* untuk data *training*, *testing*, dan *holdout* adalah 72,4%, 72,7%, dan 73,5%. Sedangkan untuk prediksi *true true positive* untuk data *training*, *testing*, dan *holdout* adalah 61,5%, 64,6%, dan 60,7%.

Selain tingkat prediksi, hasil analisis menunjukkan bahwa *jumlah pekerja* adalah karakteristik utama dari pelaku usaha yang memiliki bobot paling tinggi dalam memprediksi status keikutsertaan pemanfaatan insentif pajak, diikuti oleh besaran omzet tahunan dan pangsa pasar utama dari pelaku usaha. Untuk membahas lebih detail bagaimana ketiga variabel ini berhubungan dengan status pemanfaatan insentif pajak sehingga secara intuitif lebih mudah dipahami, penulis akan membahas secara deskriptif.²¹

²¹ Meski menyediakan informasi mengenai urutan variabel yang paling penting, hasil analisis ANN belum menyediakan informasi mengenai beberapa hal penting, misalnya: (i) bagaimana hubungan antara hubungan antara jumlah pekerja pelaku usaha dengan tingkat partisipasi pelaku usaha dalam insentif pajak? Apakah semakin besar jumlah pekerja maka kecenderungan pelaku insentif juga akan semakin tinggi? (ii) bagaimana hubungan antara tingkat omzet pelaku usaha dengan tingkat partisipasi pelaku usaha dalam insentif pajak? Apakah sifat hubungan juga positif? dan (iii) bagaimana hubungan antara pangsa pasar utama status pelaku usaha dengan tingkat partisipasi pelaku usaha dalam insentif pajak?

Data terkait *jumlah pekerja, omzet tahunan dan pangsa pasar utama* pelaku usaha dapat dikelompokkan menjadi dua klaster berdasarkan status pemanfaatan insentif.²² Selanjutnya data ini dibandingkan secara proporsional berdasarkan kategori yang ada.²³ Untuk memudahkan interpretasi, perbandingan secara proporsional dari jumlah pekerja, omzet tahunan, dan pangsa pasar utama dari pelaku berdasarkan status pemanfaatan insentif disajikan secara visual sebagaimana terlihat dalam Grafik 5.

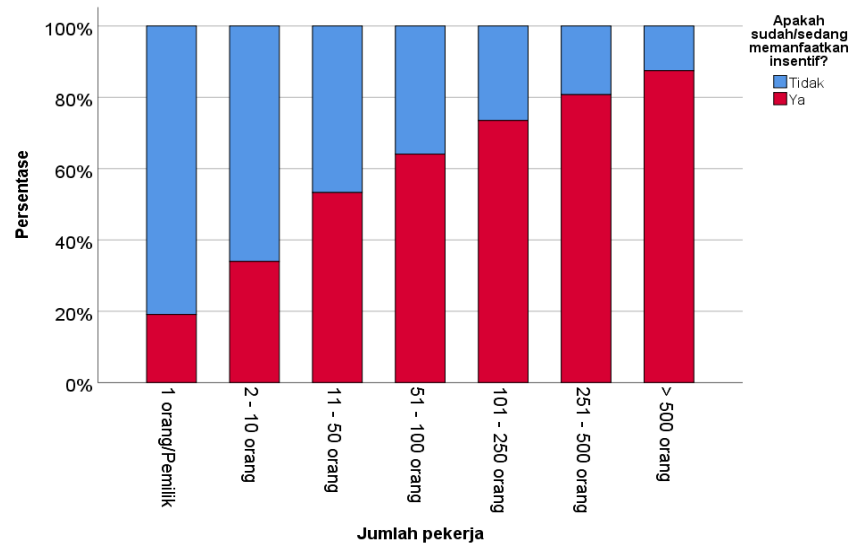
Grafik 5.A menunjukkan secara visual proporsi jumlah pelaku usaha yang ikut dan tidak ikut insentif (sumbu *y*) berdasarkan kelompok jumlah pekerja dari pelaku usaha (sumbu *x*). Grafik ini menunjukkan bahwa semakin besar jumlah pekerja, kecenderungan pelaku usaha tersebut ikut insentif akan semakin tinggi. Misalnya, di bagian kiri dari grafik ini terlihat bahwa di kelompok pelaku usaha dengan jumlah pekerja antara 2 s.d. 10 orang memiliki proporsi 67% pelaku usaha tidak ikut insentif dan 33% pelaku usaha ikut insentif. Proporsi jumlah pelaku usaha yang tidak ikut insentif ini semakin menurun, seiring dengan besarnya jumlah pekerja yang dimiliki pelaku usaha. Misalnya, di kelompok pelaku usaha dengan pekerja di atas 500 orang, proporsi pelaku usaha yang tidak ikut insentif berkurang signifikan menjadi sebesar 13% dan proporsi pelaku usaha yang ikut insentif naik menjadi 87%.

²² Data dalam studi ini memiliki porsi yang relatif seimbang untuk pelaku usaha yang ikut insentif (49%) dan yang tidak ikut insentif (51%).

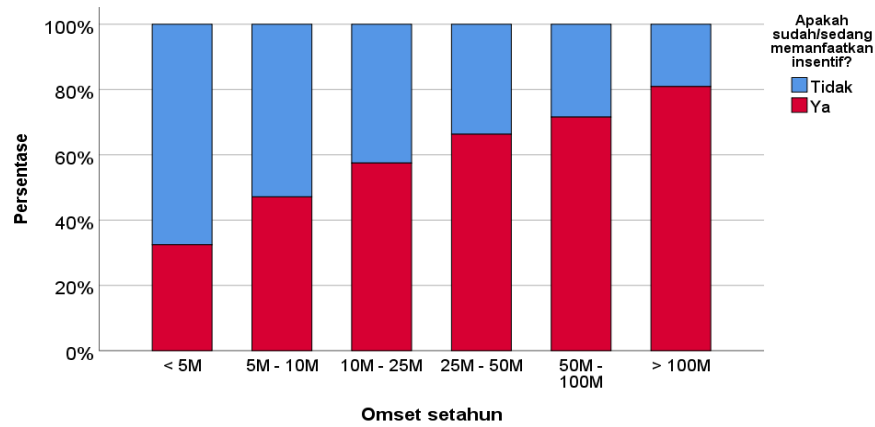
²³ Misalnya, terdapat 4.092 pelaku usaha dengan jumlah pekerja 2-10 orang dengan jumlah pemanfaat insentif dan non-pemanfaat insentif adalah masing-masing 1.368 dan 2.724. Proporsi untuk kelompok ini adalah 67% untuk pelaku usaha pemanfaat insentif (1.368 : 4.092) dan 33% untuk non-pemanfaat insentif (2.724 : 4.092).

Grafik 5: Proporsi Keikutsertaan Pelaku Usaha (n=12.361)

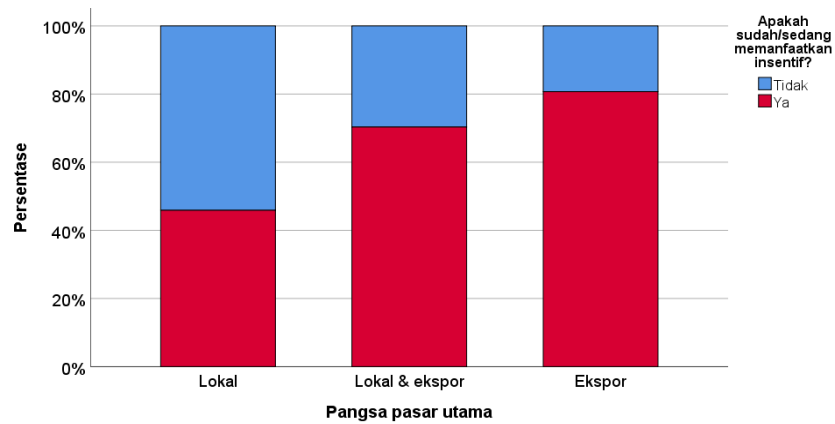
Panel 5. A



Panel 5. B



Panel 5. C



Catatan: Grafik ini menunjukkan proporsi jumlah pelaku usaha pemanfaat insentif pajak dan non-pemanfaat insentif pajak berdasarkan jumlah pekerja (panel 5.A), besaran omzet (panel 5.B), dan pangsa pasar utama (panel 5.C). Dalam panel 5.A terlihat bahwa proporsi pelaku usaha pemanfaat insentif semakin tinggi seiring semakin besarnya jumlah pekerja. Dalam panel 5.B terlihat bahwa kenaikan omzet tahunan juga beriringan dengan kenaikan proporsi pelaku usaha pemanfaat insentif. Untuk panel 5.C, terlihat pelaku usaha dengan pangsa pasar utama ekspor memiliki proporsi pelaku usaha pemanfaat insentif yang tertinggi, sementara pelaku usaha lokal memiliki proporsi paling kecil.

Selanjutnya dalam Grafik 5.B kita dapat melihat bahwa semakin besar omzet tahunan pelaku usaha, tingkat keikutsertaan pelaku usaha dalam insentif pajak juga cenderung meningkat. Misalnya, di bagian paling kiri dari grafik ini terlihat bahwa di kelompok pelaku usaha dengan omzet kurang dari Rp 5 miliar memiliki proporsi 68% pelaku usaha yang tidak ikut insentif dan 32% pelaku usaha ikut insentif. Proporsi jumlah pelaku usaha yang tidak ikut insentif ini semakin menurun, seiring dengan naiknya besaran kelompok omzet. Misalnya, di kelompok pelaku usaha dengan omzet tahunan di atas Rp 100 miliar, proporsi pelaku usaha yang tidak ikut insentif berkurang signifikan menjadi sebesar 19% dan proporsi pelaku usaha yang ikut insentif 81%.

Selanjutnya, dalam Grafik 5.C menunjukkan proporsi jumlah pelaku usaha yang ikut insentif dan tidak ikut insentif berdasarkan pangsa pasar utama pelaku usaha. Meski hubungan korelasional tidak dapat disimpulkan karena sifat nominal dari variabel ini, kita dapat melihat perbedaan proporsi dari ketiga kelompok pelaku usaha yang ada di grafik ini. Dalam hal ini, pelaku usaha dengan pangsa pasar utama lokal memiliki proporsi pelaku usaha non-pemanfaat insentif yang lebih besar (55%) dibandingkan pelaku usaha dengan pangsa pasar utama ekspor (70%).

Hal lain yang menarik dari hasil studi ini adalah konsistensi variabel *jumlah pekerja* dalam *robustness tests* berdasarkan lokasi pelaku usaha. Hasil uji menunjukkan bahwa variabel ini memiliki nilai IVI yang paling besar di seluruh kelompok pulau (lihat kembali Tabel 7). Artinya, prediktor paling kuat keikutsertaan pelaku usaha di pulau Jawa, Sumatra, Kalimantan, Sulawesi, Bali – Nusa Tenggara, dan Papua – Maluku adalah jumlah pekerja yang dimiliki pelaku usaha. Sementara itu, besaran omzet

tahunan sebagai variabel terpenting kedua berdasarkan hasil *robustness tests* menunjukkan hasil yang sedikit berbeda. Variabel ini merupakan variabel terpenting kedua di kelompok pulau Jawa, Sumatra, Sulawesi, dan Papua – Maluku. Untuk kelompok pulau Kalimantan dan Bali – Nusa Tenggara, variabel terpenting kedua adalah pangsa pasar utama dari pelaku usaha. Di dua kelompok pulau ini, omzet tahunan pelaku usaha merupakan variabel terpenting ketiga.

Hasil *robustness tests* juga menunjukkan hasil yang sedikit berbeda untuk variabel pangsa pasar utama. Variabel ini menduduki posisi ketiga di kelompok pulau Sumatra dan Sulawesi.²⁴ Yang menarik, variabel terpenting ketiga di kelompok pulau Jawa dan Papua – Maluku bukan pangsa pasar utama dari pelaku usaha, melainkan sifat usaha utama dari pelaku usaha (yaitu produsen atau non-produsen).

Informasi yang dihasilkan dari hasil analisis ini relevan dalam pengambilan keputusan yang bersifat strategis—misalnya hasil analisis ini dapat digunakan untuk membuat segmentasi kelompok pelaku usaha yang dianggap perlu untuk dinaikkan tingkat keikutsertaannya. Misalnya, otoritas pajak atau yang terkait dapat menggunakan pendekatan-pendekatan ekonomi perilaku (*behavioural economics*) untuk secara tersegmentasi melakukan intervensi kebijakan.²⁵ Meski demikian, hasil analisis ini

²⁴ Perlu dicatat, di kelompok pulau Kalimantan dan Bali – Nusa Tenggara, variabel ini merupakan terpenting kedua, bertukar posisi dengan variabel *omzet setahun* yang berada di posisi ketiga.

²⁵ Sebagai ilustrasi, untuk pelaku usaha yang belum memanfaatkan insentif dapat disampaikan pesan yang *customised*. Misalnya, untuk pelaku usaha dengan jumlah pekerja yang banyak atau memiliki omzet tahunan yang besar, namun belum memanfaatkan insentif pajak, maka pesan yang disampaikan dapat menggunakan pendekatan *peer-pressure* dengan merujuk pada hasil analisis ini. Misalnya, untuk pelaku usaha dengan jumlah pekerja antara 251 s.d. 500 orang yang belum ikut insentif pajak, narasi pesan yang dapat disampaikan: “8 dari 10 pelaku usaha yang memiliki jumlah pekerja seperti perusahaan Anda sudah memanfaatkan insentif fiskal. ...” Diskusi lebih lanjut mengenai penggunaan ekonomi perilaku di area perpajakan, misalnya, dapat dilihat dalam Weber et al. (2014). Contoh penggunaan pendekatan *behavioural economics* untuk melakukan intervensi

tidak dapat menjelaskan lebih spesifik mengenai penyebab beberapa variasi dari hasil temuan ini. Selain karena di luar cakupan dari tujuan studi, hal ini juga disebabkan karena, misalnya, status ‘pemanfaatan insentif pajak’ dalam studi ini bersifat umum; dalam arti tidak dibedakan secara spesifik jenis insentif pajak apa yang dimanfaatkan atau berapa jumlah total insentif yang dimanfaatkan.

4 SIMPULAN DAN SARAN

4.1 Simpulan

Studi ini memiliki tiga tujuan. Tujuan pertama adalah mengetahui seberapa akurat pendekatan ANN dapat memprediksi keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak. Untuk mendapatkan jawaban, studi ini menggunakan modul MLP dari ANN dan membangun model dengan formasi 60%-20%-20%. Hasil prediksi dengan modul MLP memiliki tingkat akurasi hampir 70% untuk data *testing* maupun *holdout*. Hasil studi ini menunjukkan bahwa ANN dapat digunakan untuk memprediksi apakah pelaku usaha akan memanfaatkan insentif pajak atau tidak berdasarkan karakteristik tertentu yang dimilikinya dengan tingkat akurasi yang cukup bagus. Meski demikian, perlu dicatat bahwa prediksi yang dihasilkan lebih akurat untuk kelompok pelaku usaha yang tidak ikut memanfaatkan (*true negative*)—yang berada di kisaran 73% sampai dengan 74%. Prediksi untuk kelompok pelaku usaha yang ikut insentif (*true positive*) berada di 61% sampai dengan 65%.

perilaku di Indonesia, misalnya, dapat dilihat di <https://www.bi.team/publications/encouraging-earlier-tax-returns-in-indonesia/> (diakses 10 Agustus 2021).

Tujuan kedua dari studi ini adalah menentukan tiga faktor yang paling berpengaruh terhadap tingkat keikutsertaan pelaku dalam pemanfaatan insentif perpajakan. Hasil studi ini menunjukkan bahwa faktor pertama yang paling menentukan keikutsertaan pelaku usaha adalah jumlah pekerja yang dimiliki. Variabel *jumlah pekerja* memiliki nilai *normalised IVI* tertinggi. Hasil analisis menunjukkan bahwa semakin banyak jumlah pekerja yang dimiliki pelaku usaha, semakin besar kemungkinan pelaku usaha tersebut memanfaatkan insentif pajak. Faktor terpenting kedua yang menentukan keikutsertaan pelaku usaha adalah nilai omzet tahunan. Dalam hal ini, semakin tinggi omzet pelaku usaha, semakin besar kemungkinan pelaku usaha tersebut untuk memanfaatkan insentif pajak. Penting untuk dicatat bahwa terdapat hubungan korelasional positif antara jumlah pekerja dengan omzet, meskipun hubungan ini tidak kuat (*moderate*). Selanjutnya, pangsa pasar utama dari pelaku usaha adalah faktor terpenting ketiga yang menentukan tingkat pemanfaatan insentif pajak. Pelaku usaha dengan pangsa pasar utama ekspor memiliki kecenderungan yang paling tinggi dalam pemanfaatan insentif pajak, disusul oleh pelaku usaha dengan pangsa pasar utama campuran antara ekspor dan lokal.

Tujuan ketiga studi ini adalah untuk mengetahui apakah faktor-faktor penentu yang teridentifikasi dalam model ANN bersifat *robust* ketika dilakukan *sampling variability*. *Sampling variability* dilakukan dengan mereplikasi model ANN dengan data dari enam kelompok pulau. Hasil analisis menunjukkan bahwa jumlah pekerja yang dimiliki pelaku usaha konsisten memiliki nilai *normalised IVI* tertinggi di seluruh kelompok pulau (Jawa, Sumatra, Kalimantan, Sulawesi, Bali – Nusa Tenggara, dan Papua –

Maluku). Hasil analisis juga menunjukkan bahwa besarnya omzet tahunan pelaku usaha juga cukup konsisten sebagai faktor terpenting kedua. Nilai normalised IVI dari variabel *omzet setahun* konsisten berada di urutan kedua di empat kelompok pulau (Jawa, Sumatra, Sulawesi, dan Papua – Maluku). Hal ini menunjukkan bahwa karakteristik utama dari pelaku usaha yang teridentifikasi memiliki hubungan kuat dengan keikutsertaan pelaku usaha dalam insentif pajak di enam kelompok pulau (Jawa, Sumatra, Kalimantan, Sulawesi, Bali – Nusa Tenggara, dan Papua – Maluku) bersifat homogen.

Dapat disimpulkan bahwa, dengan pendekatan ANN, keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak selama masa pandemi Covid-19 dapat diprediksi dengan tingkat akurasi yang cukup memadai. Informasi yang dihasilkan juga relevan dalam pengambilan keputusan yang bersifat strategis—misalnya hasil analisis ini dapat digunakan untuk membuat segmentasi kelompok pelaku usaha yang dianggap perlu untuk dinaikkan tingkat keikutsertaannya. Meski demikian, studi ini tidak dapat memberikan informasi lebih jauh mengenai, misalnya, mengapa—atau bagaimana—faktor seperti jumlah pekerja, nilai omzet, atau pangsa pasar utama merupakan faktor terpenting dalam menentukan keikutsertaan pelaku usaha dalam pemanfaatan insentif pajak. Untuk mengetahui hal tersebut diperlukan data dan strategi empiris yang berbeda dengan studi ini.

4.2 Saran

Studi lebih lanjut mengenai penggunaan pendekatan ANN untuk memprediksi keikutsertaan pelaku usaha dalam memanfaatkan insentif dapat dilakukan

menggunakan data administratif yang ada di dalam sistem informasi Direktorat Jenderal Pajak. Selain penggunaan data administratif, analisis yang lebih spesifik dan mendalam juga dapat dilakukan, misalnya, berdasarkan per jenis insentif—misalnya khusus untuk insentif PPh Pasal 21 DTP, insentif pembebasan PPh Pasal 22 impor, insentif pengurangan angsuran PPh Pasal 25, atau insentif pengembalian pendahuluan PPN dipercepat—atau berdasarkan sektor. Terdapat kemungkinan yang cukup besar bahwa hasil analisis lanjutan ini akan menghasilkan temuan-temuan baru yang menarik sekaligus relevan dengan proses pembuatan dan implementasi kebijakan. Selain itu, studi lebih lanjut juga dapat memasukkan variabel atau karakteristik lain dari pelaku usaha ke dalam model ANN, misalnya umur usaha, tingkat profitabilitas, nilai aset, struktur modal, atau bentuk badan usaha. Terakhir, untuk mendapatkan gambaran yang lebih lengkap sekaligus lebih spesifik, studi lebih lanjut dapat melengkapi hasil analisis dengan pendekatan lain—misalnya menggunakan *Decision Tree* atau *Cluster Analysis*.

5 REFERENSI

- Aryadoust, V & Baghaei, P 2016, 'Does EFL readers' lexical and grammatical knowledge predict their reading ability? insights from a perceptron artificial neural network study', *Educational Assessment*, vol. 21, no. 2, pp. 135-156.
- Babbie, ER 2010, *The practice of social research*, 12th edn, Wadsworth, London.
- Bahrammirzaee, A 2010, 'A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems', *Neural Computing and Applications*, vol. 19, no. 8, pp. 1165-1195.
- BPS (Badan Pusat Statistik) 2020a, *Laju Pertumbuhan PDB menurut Pengeluaran (Persen) 2020*, Badan Pusat Statistik, diakses 22 Juli 2021, <<https://www.bps.go.id/indicator/169/108/2/-seri-2010-laju-pertumbuhan-pdb-menurut-pengeluaran.html>>.
- BPS (Badan Pusat Statistik) 2020b, *Analisis Hasil Survei Dampak Covid-19 Terhadap Pelaku Usaha*, Badan Pusat Statistik, diakses 21 Juli 2021,

<<https://www.bps.go.id/publication/2020/09/15/9efe2fbda7d674c09ffd0978/analisis-hasil-survei-dampak-covid-19-terhadap-pelaku-usaha.html>>.

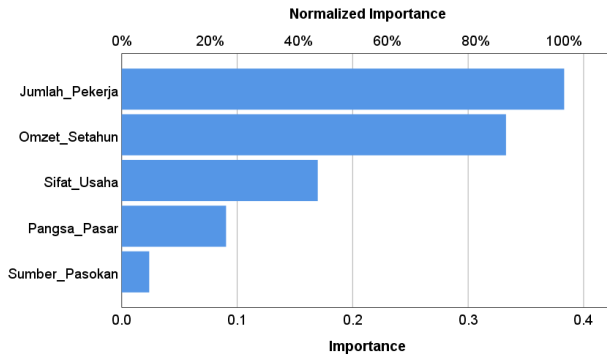
- Bekesiene, S, Smaliukiene, R & Vaicaitiene, R 2021, 'Using artificial neural networks in predicting the level of stress among military conscripts', *Mathematics*, vol. 9, no. 6, article 626.
- Chen, J-H, Su, M-C, Chen, C-Y, Hsu, F-H & Wu, C-C 2011, 'Application of neural networks for detecting erroneous tax reports from construction companies', *Automation in Construction*, vol. 20, no. 7, pp. 935-939.
- Denton, JW, Sayeed, L, Perkins, ND & Moorman, AH 1995, 'Neural networks to classify employees for tax purposes', *Accounting, Management & Information Technology*, vol. 5, no. 2, pp. 123-138.
- DJP (Direktorat Jenderal Pajak) 2021, *Buku statistik perpajakan 2020*, Direktorat Jenderal Pajak, Jakarta.
- Fawcett, T 2006, 'An introduction to ROC analysis', *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874.
- Haykin, SS 1999, *Neural networks: a comprehensive foundation*, 2nd edn, Macmillan College Publishing, New York.
- IBM (International Business Machines) 2017, *IBM SPSS Decision Trees 25*, IBM Corporation, diakses 15 Juni 2021, ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/25.0/en/client/Manuals/IBM_SPSS_Decision_Trees.pdf.
- IBM (International Business Machines) 2019, *IBM SPSS Neural Networks 26*, IBM Corporation, diakses 21 Juli 2021, <https://www.ibm.com/docs/SSLVMB_26.0.0/pdf/en/IBM_SPSS_Neural_Network.pdf>.
- Jang, SB 2019, 'A design of a tax prediction system based on artificial neural network', paper yang dipresentasikan pada International Conference on Platform Technology and Service (PlatCon), Jeju, 28-30 Januari.
- Jupri, M & Sarno, R 2018, 'Taxpayer compliance classification using C4.5, SVM, KNN, Naive Bayes and MLP', paper yang dipresentasikan pada International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, 6-7 Maret.
- Kementerian Keuangan RI 2021, *Laporan Keuangan Pemerintah Pusat Tahun 2020 (Audited)*, Kementerian Keuangan RI, diakses 21 Juli 2021, <<https://www.kemenkeu.go.id/media/18103/lkpp-2020.pdf>>.
- Lin, C-H, Lin, I-C, Wu, C-H, Yang, Y-C & Roan, J 2012, 'The application of decision tree and artificial neural network to income tax audit: the examples of profit-seeking enterprise income tax and individual income tax in Taiwan', *Journal of the Chinese Institute of Engineers*, vol. 35, no. 4, pp. 401-411.

- Neuman, WL 2011, *Social research methods: qualitative and quantitative approaches*, 7th edn, Allyn & Bacon, Massachusetts.
- OECD (Organisation for Economic Co-operation and Development) 2020, *Tax and Fiscal Policy in Response to the Coronavirus Crisis: Strengthening Confidence and Resilience*, Organisation for Economic Co-operation and Development, diakses 21 Juli 2021, <<http://www.oecd.org/coronavirus/en/>>.
- Pérez López, C, Delgado Rodríguez, MJ & Lucas Santos, S 2019, 'Tax fraud detection through neural networks: an application using a sample of personal income taxpayers', *Future Internet*, vol. 11, no. 4, article 86.
- Ripley, BD 1996, *Pattern recognition and neural networks*, Cambridge University Press, Cambridge.
- Sánchez-Serrano, JR, Alaminos, D, Garcia-Lagos, F & Callejón-Gil, AM 2020, 'Predicting audit opinion in consolidated financial statements with artificial neural networks', *Mathematics*, vol. 8, no. 8, article 1288.
- Schober, P, Boer, C & Schwarte, LA 2018, 'Correlation coefficients: appropriate use and interpretation', *Anesthesia & Analgesia*, vol. 126, no. 5, pp. 1763-1768.
- Weber, TO, Fooker, J & Herrmann, B 2014, 'Behavioural economics and taxation', European Commission Taxation Papers Working Paper, no. 41, European Commission, Luxembourg, diakses 21 Juli 2021, <http://knjiznica.sabor.hr/pdf/E_publicacije/Behavioural_economics_and_taxation.pdf>.
- Zhang, G, Hu, MY, Patuwo, BE & Indro, DC 1999, 'Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis', *European Journal of Operational Research*, vol. 116, no. 1, pp. 16-32.

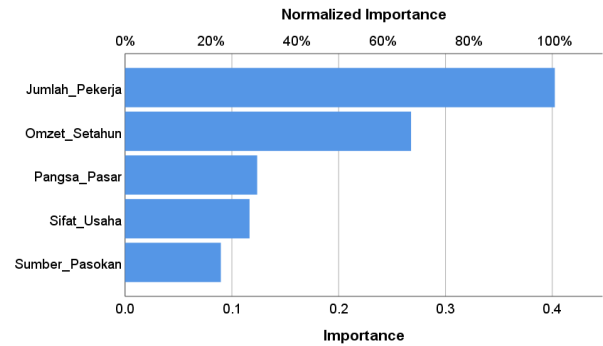
LAMPIRAN

Hasil *robustness test* di enam kelompok pulau: *Independent Variable Importance*

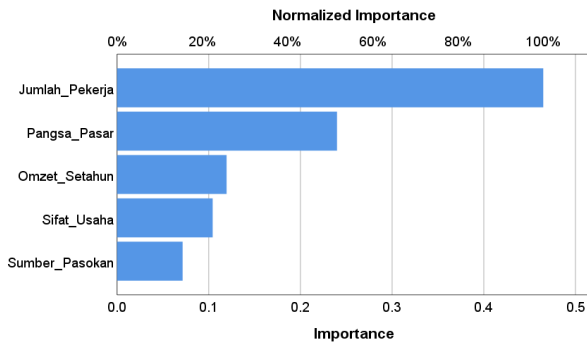
Jawa (n = 8.681)



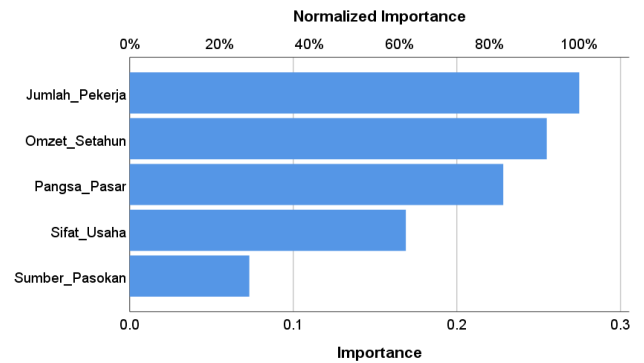
Sumatra (n = 1.633)



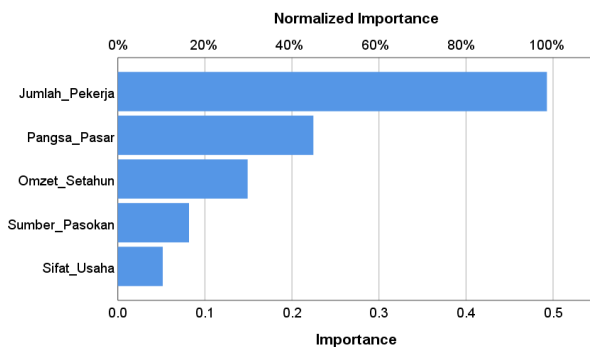
Kalimantan (n = 706)



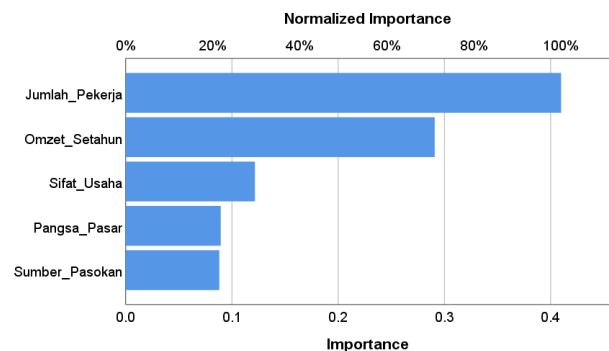
Sulawesi (n = 612)



Bali – Nusa Tenggara (n = 573)



Papua – Maluku (n = 156)



Keterangan: Grafik di atas merupakan hasil analisis ANN secara terpisah menggunakan data pelaku usaha dari enam pulau utama di Indonesia: Jawa, Sumatra, Kalimantan, Sulawesi, Bali – Nusa Tenggara, dan Papua – Maluku. Nilai IVI menunjukkan bahwa variabel *jumlah pekerja* konsisten memiliki urutan tertinggi sebagai variabel independen yang memiliki tingkat kepentingan paling besar dalam prediksi ANN di enam kelompok pulau. Sedangkan *omzet setahun* merupakan variabel independen terpenting kedua untuk kelompok pulau Jawa, Sumatra, Sulawesi, dan Papua – Maluku.