energies

Article

Artificial Neural Networks as a Tool to Understand Complex Energy Poverty Relationships: The Case of Greece

Lefkothea Papada and Dimitris Kaliampakos *

School of Mining and Metallurgical Engineering, National Technical University of Athens, Zografou Campus, 15772 Zografou, Greece; lefkipap@metal.ntua.gr

* Correspondence: dkal@central.ntua.gr

Abstract: The present paper provides an innovative approach in the existing methods of studying energy poverty, i.e., a crucial socio-economic challenge of the past decade in Europe. Since the literature has shown that conventional statistical models lack effectiveness in handling unconventional relationships between variables and present limitations in terms of accurate classification and prediction, the paper explores the ability of Artificial Intelligence and, particularly, of Artificial Neural Networks (ANNs), to successfully predict energy poverty in Greece. The analysis included the prediction of seven energy poverty indicators (output indicators) based on certain socio-economic/geographical factors (input variables), via training an ANN, i.e., the Multilayer Perceptron. Three models (Model A, Model B and Model C) of different combinations of the input variables were tested for each one of the seven indicators. The analysis showed that ANNs managed to predict energy poverty at a remarkably good level of accuracy, ranging from 61.71% (lowest value) up to 82.72% (highest accuracy score). The strong relationships that came up on the examined cases confirmed that ANNs are a promising tool towards a deeper understanding of the energy poverty roots, which in turn can lead to more targeted policies.

Keywords: energy poverty; Artificial Intelligence; Artificial Neural Networks; indicators; Greece

check for updates

Citation: Papada, L.; Kaliampakos, D. Artificial Neural Networks as a Tool to Understand Complex Energy Poverty Relationships: The Case of Greece. *Energies* **2024**, *17*, 3163. https://doi.org/10.3390/en17133163

Academic Editors: David Borge-Diez, Pedro L. Lomas and Yu Hao

Received: 27 May 2024
Revised: 15 June 2024
Accepted: 25 June 2024
Published: 27 June 2024

1. Introduction

1.1. The Problem of Energy Poverty

Energy poverty is a crucial socio-economic challenge, leading to the deprivation of individuals from fundamental standards of life quality. The lack of access to electricity and to clean and modern energy services still remains a luxury for masses of people (760 million for electricity and 2.3 billion people for clean energy services) living in developing countries [1]. However, in the developed world, the concept of energy poverty is primarily associated with the affordability of energy, indicating the difficulty or inability of individuals to adequately address their energy needs at a reasonable price.

This issue has come to the fore within policy issues so much in Europe and also in Greece over the past decade, as a result of a financial crisis at first, but it has worsened even further in the last couple of years due to global geopolitical events (e.g., war in Ukraine), which have caused great instability and high increases in energy prices. In this context, large parts of the population spend large amounts of money on energy expenses compared to their income, struggle to maintain sufficient heating levels within their residence, face energy bill arrears or damp/mold problems in their residence, reduce energy consumption to unacceptable levels, and generally experience this difficulty in so many different ways.

Especially, Greece is placed on top of countries facing severe energy poverty issues [2]. The prolonged financial crisis over the past 14 years has ranked Greece fifth among the European countries at risk of poverty or social exclusion, affecting 26.1% of its population [3]. During the past decade, there has been a remarkable rise in fuel costs, combined with a considerable drop in household incomes, while in the last couple of years, a dramatic increase in food prices has also been noted. Overall, this situation has appeared in alignment with broader global conditions, including the climate crisis, COVID-19 pandemic, refugee crisis, and geopolitical conditions of instability. It should also be noted that the energy sector has witnessed significant changes in Greece, such as the liberalization of the energy market in line with the "energy transition" of the European Union, affecting the energy production, transmission, and distribution system. These alterations have operated within a speculative environment, playing a main role in the escalation of energy prices [4]. Another evident indication of energy poverty in the country over the last years has been the extensive use of inappropriate and cheap materials as fuels for heating, particularly in cases of open fireplaces and stoves, leading to environmental and health issues.

Energy poverty research in Greece has gained particular attention in recent years, mainly due to the rapid deterioration of the phenomenon. Indicatively, Atsalis et al. [5] studied the 10% objective indicator, which represents energy expenses surpassing 10% of the household's income, using official data provided by the Hellenic Statistical Authority (Household Budget Survey). The results showed that in 2013, around 20–25% of households in Greece were energy poor, showing a notable rise from the shares of 9–13% in 2008. Additionally, a thorough investigation of the energy poverty problem in the whole country (Greece) was carried out by Papada and Kaliampakos [6] via a primary survey, concluding that 58% of Greek households encountered energy poverty in 2015, as determined by the 10% indicator (actual expenses as a share of disposable household income). Furthermore, Papada and Kaliampakos [7] implemented a similar primary survey focusing on the mountainous regions of Greece, highlighting the severity of the problem in these areas. The results demonstrated that a significant share (73.5%) of households living in mountainous Greece faced energy poverty in 2015 due to the combined impact of high energy expenses and reduced incomes.

Other studies have focused on certain geographical areas, highlighting the intensity of the problem in different regions, such as Thessaloniki Urban Complex [8], Western and Central Macedonia [9], the Attica region [10] and Athens [4,11]. Furthermore, new indices have been suggested, such as, for example, the "Degree of Coverage of Energy Needs" (DCEN) [12], revealing that almost half of all households in Greece compress their energy needs. Lyra et al. [13] used logistic regression models based on data from Eurostat (EU-SILC survey), demonstrating that 40% of Greek households encounter energy poverty issues, while also certain factors were identified as determinants of the problem, i.e., dwelling type, household income, educational level, and location of residence. Similarly, Kalfountzou et al. [14] used binary logistic regression models in order to predict different indicators (10% indicator, 2M, M/2) based on socio-economic variables. Finally, Halkos and Kostakis [15] showed that about 9–10% tend to be steadily affected by energy poverty, with variables like housing characteristics, income level, educational level, employment status, and migration background impacting the probability of encountering this issue.

It is noteworthy that the majority of previous research used conventional statistical tools for their analysis, e.g., logistic regression models. However, most traditional statistical models assume linearity or quadratic relationships between explanatory and dependent variables, thereby lacking effectiveness in handling unconventional relationships and, hence, present limitations in terms of accurate classification and prediction of dependent variables [16,17]. On the other hand, prediction models based on machine learning offer flexibility in handling unconventional relationships between variables and display enhanced predictive capability [17,18]. Yet, machine learning techniques also pose constraints due to the complexity of understanding the prediction process and the challenge of verifying the statistical significance of independent variables. Despite these constraints and given the increasing significance of the fast and focused identification of policy targets, there is a growing need for prediction models with enhanced predictive power [17], an outcome that has been also reached for the field of energy poverty alleviation [19]. Hence, machine learning algorithms started attracting more and more interest in forecasting energy poverty.

1.2. Artificial Neural Networks and Energy Poverty

The inception of Artificial Neural Networks (ANNs) originated from an attempt to comprehend the way the human brain functions and mimic its cognitive capabilities. The challenge of ANNs is to be able to make decisions and take actions under conditions of uncertainty, or even handle conditions with little previous experience [20]. ANNs have demonstrated a noteworthy ability in addressing multi-variable systems, non-linearities, and uncertainties, often surpassing alternative methods in terms of precision and adaptability [21]. Especially the ability of ANNs to handle vast amounts of data from diverse sources, as well as to extract practical insights, has been reported as the one that distinguishes them [22–24]. In terms of their weak points, ANNs need a large amount of training data, display sensitivity to input data quality as well as a greater computational burden, are prone to overfitting, and are characterized by a "black box" operation that can hinder interpretability and transparency in decision-making [25].

Specifically, ANNs are mathematical models comprising interconnected processing nodes (neurons), organized in a predetermined structure (layers) [26]. Their primary characteristic is the ability to conduct extensive parallel processing of input data, as opposed to conventional mathematical models that rely on a sequential execution of mathematical functions [27]. Their operation relies on data collection in a cause–effect format (input–output), which are inserted into the network for training. After this process, the neural network is able to detect the underlying relationships within the data, while adjusting connection weights accordingly [20]. Through this iterative process, the network is educated and developed through practice and, hence, it is able to provide estimations for uncertain circumstances or incomplete datasets [28].

According to the literature and, before focusing on energy poverty, various references of ANNs used for predicting energy features, especially energy consumption, have been detected. Indicatively, Rajic et al. [29] introduced a model for predicting the energy consumption (electricity and heating consumption) of Serbian households, based on 15 macroeconomic input parameters such as population, transportation, unemployment rate, average wage, industrial activities, etc., using a neural network in Matlab, referred to as the "Endocrine NARX neural network" (ENARX). Yan Cheng-wen and Jian [30] used a backpropagation neural network to predict the energy consumption of buildings at different climate zones. Pavićević and Popović [31] developed models using various neural network architectures to forecast the electricity price and the electricity load in Montenegro. Heghedus et al. [32] used four different types of neural networks for electricity consumption forecasting, while Tardioli et al. [33] used three models, i.e., generalized linear models, neural networks, and support vector machines, in order to analyze the demand profiles of numerous buildings sharing the same typology.

The literature shows that there is a large number of works that use Artificial Intelligence to study partial aspects of energy poverty or areas closely linked to the problem [34]. Yet, there is a limited quantity of works that use Artificial Intelligence to solely study energy poverty. In fact, machine learning techniques—as well as ANNs specifically—focusing exclusively on energy poverty have begun to attract attention mainly during the last decade. Indicatively, Longa et al. [35] employed a decision tree classifier (XGBoost) to predict the likelihood of energy poverty (household energy expenditure vs. disposable household income = four risk categories) in the Netherlands, based on income and various socio-economic variables (inputs), like property status, economic value of the property, household size, age of the house, and average population density. Notably, they highlighted that machine learning could serve as a valuable tool for monitoring energy poverty, potentially helping in designing appropriate policy interventions. Al Kez et al. [36] employed a random forest model to predict the "LILEE indicator" (Low Income Low Energy Efficiency) of energy poverty—four risk categories—in the UK, based on satellite remote sensing and socioeconomic data, using data from the UK English Housing Survey (EHS).

Pino-Mejías et al. [18] applied two methods, i.e., multiple linear regression and ANNs to predict an energy poverty indicator, called the "Fuel Poverty Potential Risk Index" (FP-

PRI) in Chile/Bio-Bio Region, using data from simulations and mathematical frameworks. Following this work, Bienvenido-Huertas et al. [37] also used ANNs to predict the same output indicator (FPPRI) in diverse climate zones of Chile, or in three cities with the highest population in Chile (Santiago, Valparaiso, and Concepción). Bienvenido-Huertas et al. [38] applied three algorithms, i.e., Multilayer Perceptron (ANN), M5P, and random forest, to predict the "2M" indicator of energy poverty (high share of energy expenditure) in warm climate zones of Spain. Papada and Kaliampakos [26] used a neural network to predict energy poverty as defined by different "objective" indicators, based on "subjective" indicators. The results showed that specific human behaviors or subjective aspects can predict energy poverty at a marginally acceptable rate, approximately 56–58%. In developing countries, Abbas et al. [39] used a backpropagation neural network to predict "Multidimensional Energy Poverty" based on socioeconomic variables in Asia and Africa, using household survey data from 39 Sub-Saharan African and 20 Asian countries.

In the present paper, ANNs are used in order to predict energy poverty—as being defined by different indicators—based on certain socio-economic/geographical variables, for the case of Greece. The question investigated is whether there are potential "patterns" in the way energy poverty is related to—some or all of—the examined variables, without knowing or investigating the relationships among them. In this way, groups of people being at the highest energy poverty risk can be detected and, hence, decisions and policies regarding energy poverty alleviation can be significantly enhanced.

2. Materials and Methods

A database was primarily set up, comprising data from five different primary surveys on energy poverty in Greece. The five surveys studied similar energy poverty features on different geographical locations within the country, i.e., the first survey covering the entire country (Greece) [6], the second one studying the entire mountainous zone (mountainous regions) of Greece [7], two surveys focusing on the town of Metsovo, Greece [40,41], and a final survey concerning the settlement of Agrafa, Greece [42]. The latter two locations were selected as illustrative case studies of mountainous regions most severely affected by energy poverty in Greece. All aforementioned surveys used for the analysis were carried out between 2016 and 2020 at a household level, based on a random sampling approach. A confidence interval of 95% was used, ensuring a maximum 5% margin of error. The questionnaires of all surveys included questions encompassing a broad spectrum of topics throughout recent years, such as living and housing conditions, housing infrastructure, heating systems, personal views on energy access and well-being, quantitative data on income and energy expenses, and socio-demographic information.

The new database generated contained a total of 1754 data series, i.e., 400 instances of households residing in Greece, 400 households in mountainous regions of Greece, 643 households in the town of Metsovo, and 311 households in the settlement of Agrafa. The analysis included the prediction of certain energy poverty indicators (output indicators) based on socio-economic/geographical factors (input variables), via training an Artificial Neural Network (ANN). For this reason, seven (7) energy poverty indicators were selected as output indicators, four (4) of which are objective or expenditure-based and three (3) of which are subjective or self-reported, as described below:

1. The "10% actual" indicator, according to which a household is considered energy poor if it spends more than 10% of its disposable income on its annual energy expenses [6,43].
2. The "10% required" indicator, according to which a household is considered energy poor if it needs to spend more than 10% of its disposable income on its theoretically required annual energy expenses [44].
3. The "CEN" indicator, i.e., Compression of Energy Needs, according to which a household is considered energy poor if it spends on energy less than 80% of the amount required to cover its energy needs sufficiently [12,26].

4. The "NEPI" indicator, i.e., National Energy Poverty Index, according to which a household is considered energy poor if two conditions are simultaneously met: (i) the total annual energy cost of the household is below 80% of the amount theoretically required to meet energy needs, and (ii) the total income of the household (equivalized, based on the OECD equivalence scale) is below the poverty line, as defined in Greece, i.e., is less than 60% of the median equivalized income of all households in the country [4,45].
5. The "IW" indicator, expressing the inability to keep a home adequately warm, as also measured by Eurostat (EU Statistics on Income and Living Conditions-EU SILC survey) [46].
6. The "AB" indicator, expressing the arrears on energy bills, simulating the "Arrears on utility bills" indicator measured by Eurostat (EU SILC survey) [47], but more precisely focusing on energy expenses.
7. The "DL" indicator, expressing the problems of a leaking roof, damp walls/floors/foundation, or rot in window frames or floor, as also measured by Eurostat (EU SILC survey) [48].

The specific objective indicators were selected among a number of indicators as being considered the most critical ones, according to the authors' view. Specifically, the 10% rule is widely recognized as a key indicator of energy poverty in Europe, either measured by actual ("10% actual" indicator) or by required expenses ("10% required" indicator), focusing on energy cost, a crucial element of energy poverty [43,49]. Conversely, the "CEN" indicator addresses the overlooked yet vital factor of energy poverty, i.e., the situation in which a household either minimizes energy expenditure, or otherwise consumes less energy than necessary to fulfill energy requirements. This latter aspect is frequently disregarded in common indicators, despite being a substantial dimension of energy poverty. The "NEPI" indicator combines critical components of energy poverty, i.e., low energy cost, low energy consumption and low household income, and thus is considered a decisive indicator of energy poverty. On the other hand, subjective indicators can capture the self-reported nature of energy poverty or it can capture aspects of the problem that objective indicators may miss; hence, they are considered valuable tools for assessing the overall picture of energy poverty.

Some of the selected indicators (10% actual, 10% required, IW, AB, and DL) were already calculated in the context of the five (5) previous primary surveys [6,7,40–42], while some others (CEN and NEPI) had to be calculated from scratch. A key point of criticism associated with some of the above indicators, i.e., "10% required" indicator, "CEN" indicator, and "NEPI" indicator, is the complex calculation process needed in order to model the theoretically required energy consumption of the residence, embedded in the definitions of the above indicators. This calculation apparently depends on a number of both technical and climatic characteristics, such as the house age, the existence of thermal insulation, the climatic conditions of the region, etc. The challenge of modeling/calculating the required energy consumption has forced the majority of researchers so far to adopt the use of actual costs versus required ones in calculations, given that data on actual costs can be more easily obtained through national statistical surveys and primary surveys [26,43,49]. In the current research, this difficulty has been overcome with the use of a model/calculation tool developed by Papada and Kaliampakos [2], i.e., the "Stochastic Model of Energy Poverty" (SMEP), which enables the calculation of required energy consumption at a household level, based on both technical and climate features, as defined by the equation below:

$$\begin{aligned} &Modelled\ energy\ costs = modelled\ consumption \times\ price \\ &= E_{heat} \cdot price_{heat} + E_{cool} \cdot price_{cool} + E_{electr} \cdot price_{electr} + E_{dhw} \cdot price_{dhw} \\ &= \frac{H_{tot} \cdot HDD \cdot 24 \cdot price_{heat}}{n_h \cdot 1000} + \frac{H_{tot} \cdot CDD \cdot 24 \cdot price_{cool}}{n_c \cdot 1000} + E_{elect} \cdot price_{elect} + E_{dhw} \cdot price_{dhw} \end{aligned} \quad (1)$$

where

E_{electr}: annual energy consumption for electricity (KWh); E_{dhw}: annual energy consumption for domestic hot water (KWh); $price_{heat}$: unit price of heating (EUR/KWh); $price_{cool}$: unit price of cooling (EUR/KWh); $price_{electr}$: unit price of electricity (EUR/KWh); $price_{dhw}$: unit price of domestic hot water (EUR/KWh); H_{tot}: total heat transfer coefficient due to both convection and ventilation (W/°C); *HDD*: heating degree days (°C*days); *CDD*: cooling degree days (°C*days); n_h: performance factor of the heating energy system; n_c: performance factor of the cooling energy system.

However, the calculation of energy poverty indicators was not the target of the particular research but the means by which the new database was enriched with data of different samples, geographical areas, and time periods, so as to finally enable the application of machine learning. As a result, after the calculation process, seven (7) different datasets were set, each one including a number of socio-economic and geographical variables (house age, ownership status, number of household members, house area, and elevation) and each one of the seven (7) indicators was used each time. In other words, a distinct dataset of 1754 data series was created per indicator. The input variables and the output indicators examined are shown in Table 1.

The final selection of input variables was made upon several tests in terms of determining the most decisive factors affecting energy poverty. Notably, the geographical variable of "elevation" was taken into consideration as, on the one hand, elevation is the second most significant geographical factor—following latitude—that affects climate, and on the other hand, elevation has been proved to be the predominant factor affecting energy demand in mountainous regions [50], and Greece is a predominantly mountainous country. More precisely, mountainous areas cover over 70% of the country's territory, while they host a smaller share of the Greek population (around 30%, mountainous and semi-mountainous together) [7].

Table 1. Input variables and output indicators examined.

Input Variables	Output Indicators (One Per Time)						
House age	"10% actual" indicator	"10%" required indicator	"CEN" indicator	"NEPI" indicator	"IW" indicator	"AB" indicator	"DL" indicator
Ownership status							
Household size							
House area							
Elevation							

The core part of the analysis took place after configuring the datasets. Specifically, a machine learning tool/application called "WEKA" was used in order to train a neural network, with the aim of predicting energy poverty risk as defined by each one of the seven (7) indicators, based on different combinations of the input variables. WEKA (Waikato Environment for Knowledge Analysis), developed at the homonymous University of New Zealand and written in Java, offers a range of machine learning algorithms, along with tools for preprocessing data [51]. Multilayer Perceptron, i.e., a neural network that employs backpropagation for training, was used for the current analysis. The datasets had to be modified accordingly (detection of missing values, conversion to arff format files, etc.) so that they were compatible with the WEKA interface.

Three (3) basic models were tested for each one of the seven (7) indicators, as shown in Table 2:

- Model A: house age, ownership status and household size.
- Model B: house age, ownership status, household size and house area.
- Model C: house age, ownership status, household size, house area and elevation.

For each distinct model examined, the dataset was split into training and test sets, i.e., 70% and 30%, respectively.

Table 2. Models examined per output indicator.

Models	**Output Indicators (One Per Time)**						
Model A: House age Ownership status HH size	"10% actual" indicator	"10%" required indicator	"CEN" indicator	"NEPI" indicator	"IW" indicator	"AB" indicator	"DL" indicator
Model B: Model A + House area							
Model C: Model B + Elevation							

Various techniques were employed to enhance the performance of the neural network models. Specifically, the selection of hidden layers and nodes was adapted to each different model in order to optimize outcomes. Additionally, when the output variable classes were significantly unbalanced, the filter "SMOTE", i.e., Synthetic Minority Oversampling Technique, was applied to mitigate bias in the results. SMOTE enabled the adjustment of the relative frequency between minority and majority classes by oversampling the minority class, thus generating synthetic instances via the approach of k-nearest neighbor. Hence, numerous iterations were conducted per model tested, adjusting elements of the neural network to achieve optimal model performance.

The performance of the models was evaluated according to certain metrics. These included the "Accuracy score", indicating the correctly classified instances, as well as key performance metrics indicating the detailed accuracy by class, namely, "Precision", "Recall", "F-Measure", and "ROC Area":

- "Precision" reflects the share of correct predictions of a class within the total correct predictions of the class and incorrect predictions of the other class (sum of instances classified as a given class category of the output indicator) (Figure 1).
- "Recall" represents the share of correct predictions of a class within the total predictions of the class (correct and incorrect) (Figure 1).
- "F-Measure" combines "Precision" and "Recall" and is utilized as a general metric considering the costs of incorrect predictions.
- "ROC Area", or "Receiver Operator Characteristic Area under the curve", serves as an accuracy measure of the model, indicating the level of a random model's prediction and, ideally, aims to be the highest possible. Notably, "ROC Area" provides insights about the actual appropriateness of the neural network.

Furthermore, a 2 × 2 confusion matrix was generated for each model, delineating correct and incorrect predictions (Figure 1). In more detail, the diagonal elements of each row of the confusion matrix represent accurately predicted instances (true positives) per class category of the output indicator.

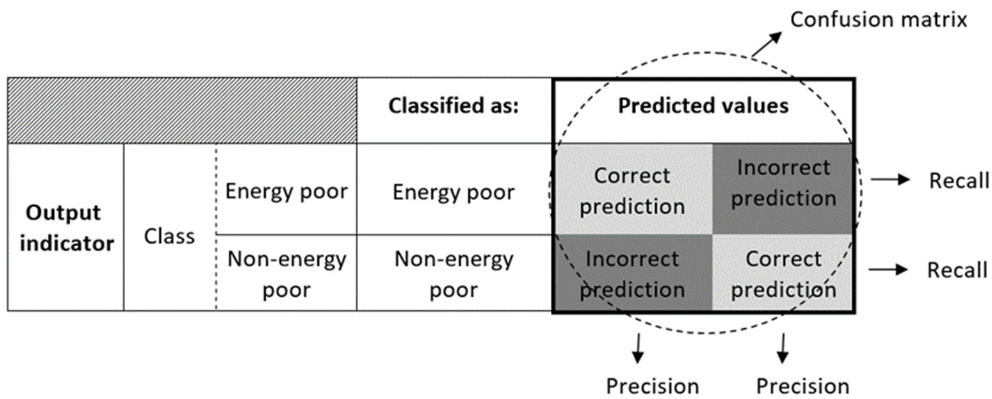

Figure 1. Main performance metrics per model (Precision, Recall and Confusion Matrix).

On the whole, the evaluation of the performance of a model entails the consideration of all the above elements collectively.

It should be reminded though that the current research does not focus on the calculation of the energy poverty rate per indicator; the aim here is to explore the ability of an ANN to accurately predict the energy poverty rate given per indicator, based on certain socio-economic and geographical variables. If this attempt is successful, then it will be possible to forecast the energy poverty risk of any population group, based only on certain input variables, with the use of ANNs.

3. Results and Discussion

Tables 3–9 present the performance of the ANN models (Model A, Model B and Model C) tested per output indicator. From the hundreds of different tests performed in terms of improving performance levels, only the best models are presented, i.e., the ones with the highest accuracy scores, combined with the most meaningful performance metrics.

3.1. Prediction of the "10% Actual" Indicator

Regarding the prediction of the "10% actual" indicator, the accuracy score ranges from 59.96% to 69.29% (Table 3). The lowest accuracy (59.96%) appears in both Models A and B, using as input variables "house age", "ownership status", "household size" (Model A), plus the "house area" (Model B). Model C, with the addition also of "elevation" in input variables, further improves the accuracy score by 9 percentage points, reaching the highest accuracy score (69.29%).

As far as the other performance metrics are concerned, "Precision", "Recall", "F-Measure", and "ROC Area" for the weighted average marginally exceed—or even did not exceed—the limit of 0.5 in both Models A and B, implying that the prediction of the examined models is random, hence the models are inappropriate. Even more, according to the diagonal parts of the confusion matrix of Models A and B, 97–98% of the energy poor households are correctly predicted but only 4–6% of the non-energy poor households are correctly predicted, indicating that the two models are unable to correctly predict the second class of the output indicator, thus they are considered non-meaningful models on the whole.

On the contrary, Model C presents a considerable improvement in the case of performance metrics, as also in confusion matrix elements. Specifically, "F-Measure" (0.693) and "ROC Area" (0.712) for the weighted average present considerably improved values, indicating that the examined model is reasonable. It is also shown that 74% of the energy poor households and 62% of the non-energy poor households are correctly predicted. Hence, contrary to the other two models, Model C manages to correctly predict both classes (energy poor and non-energy poor households) at a satisfactory level.

In other words, a neural network using all five (5) input social–geographical variables (house age, ownership status, household size, house area, and elevation) can correctly predict energy poverty—as described by the 10% indicator based on actual expenses—at a satisfactory level (69.29%).

The neural network of the best model (Model C) included one hidden layer, which consisted of seven nodes. The filter SMOTE was also applied to adjust the relative frequency between minority and majority classes, i.e., double the instances of minority classes, based on the k-nearest neighbor approach (15 nearest neighbors selected). The specific neural network is illustrated in Figure 2.

Table 3. Prediction of the "10% actual" indicator and confusion matrices (test sets presented).

	Prediction of "10% Actual" Indicator							
Input Variables	**Precision**	**Recall**	**F-Measure**	**ROC Area**	**Class**	**Accuracy**	**Confusion Matrix**	
Model A: House age Ownership status HH size	0.600	0.978	0.744	0.549	Yes	**59.96%**	**98%**	2%
	0.579	0.044	0.081	0.549	No		96%	**4%**
	0.592	0.600	0.476	0.549	(weighted avg)			
Model B: Model A + House area	0.602	0.968	0.742	0.570	Yes	**59.96%**	**97%**	3%
	0.556	0.060	0.108	0.570	No		96%	**6%**
	0.583	0.600	0.485	0.570	(weighted avg)			
Model C: Model B + Elevation	0.743	0.741	0.742	0.712	Yes	**69.29%**	**74%**	26%
	0.621	0.623	0.622	0.712	No		38%	**62%**
	0.693	0.693	0.693	0.712	(weighted avg)			

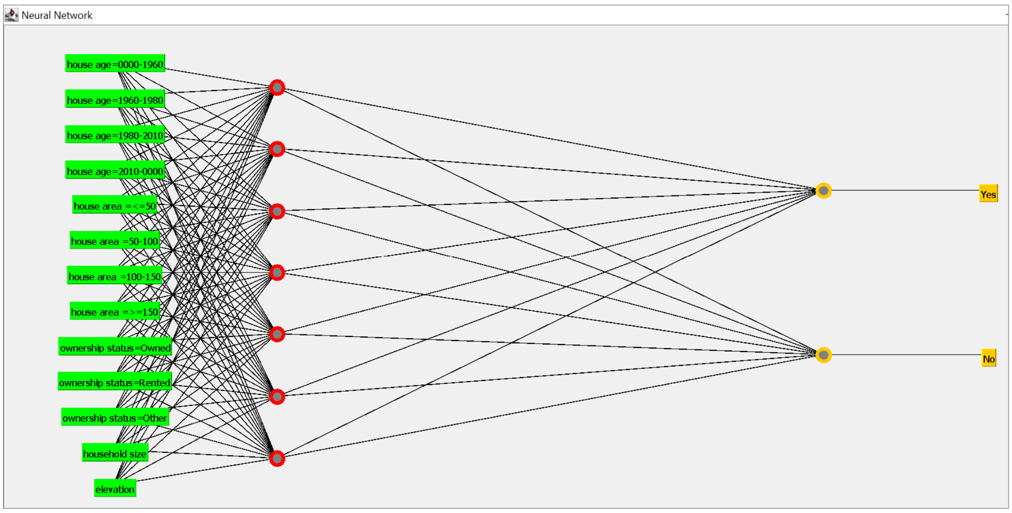

Figure 2. ANN of the model with the best performance, predicting the "10% actual" indicator.

Compared with traditional mathematical models, Kalfountzou et al. [14] used a binary logistic regression model to predict the same output variable (household energy expenditure as a share of disposable household income), using as independent variables four (4) socio-economic factors, i.e., region, household type, socio-economic situation of the household, and population density. The model presented a performance level of just 32%, which was considered the highest one among the paper's tests. As a result, ANNs demonstrate better results than traditional models in terms of energy poverty prediction, based on socio-economic variables (accuracy score of 69% vs. 32%).

Some further points of criticism should be raised here. The prediction of the "10% actual" indicator simulates the work of Longa et al. [35], who employed a decision tree classifier (XGBoost) to predict the same output variable (household energy expenditure as a share of disposable household income) in the Netherlands, based on income and five socio-economic variables (house age, ownership and value, household size, and average population density). The results showed that the model using all variables as input variables, i.e., income and the five socio-economic variables, presents an accuracy score of 80% and 73–88% accurate predictions according to the confusion matrix, whereas the model using only the five socio-economic variables as input variables presents an accuracy score of just 52% and only 31–66% accurate predictions. In other words, when income is excluded, the model turns out to be not meaningful at all, in terms of its predictive power. The main point of criticism here is that the inclusion of income—as a predictive variable—in a model predicting an output indicator that already includes the same variable (income) in its own definition is misleading and may produce biased results. The authors themselves [35] point out that income was expected to be the most influential feature in their models examined, which is a fact, indeed. Instead, the model of the present paper (Model C) built on social–geographical variables (not including income as well) and run by an ANN presents valuable metrics and a predictive power of the order of 70%, significantly higher compared to a similar model run by a decision tree classifier (69% vs. 52% and a far better confusion matrix).

Similar results about the inclusion or not of variables that are already involved in the definition of the output variable arise in Al Kez et al. [36], who employed a random forest model to predict the "LILEE indicator" (Low Income Low Energy Efficiency) in the UK. It is noteworthy that the model using income and efficiency as input variables (both already included in the definition of the output variable) plus eight additional variables presents an accuracy score of 100% (regarded as highly controversial as a prediction rate), whereas the model using only the eight variables as input variables presents an accuracy score of 67%. In other words, when income and efficiency (disputed variables) are excluded from the model, the accuracy score drops by 33%.

The above criticisms should be seriously taken into account in prediction models, so that biased results are avoided.

With regard to the predictors examined, household size and house ownership proved to be significant predictive factors in [35] for the same output indicator (10% indicator based on actual expenses), which was validated in the present paper but only combined with the rest of the social–geographical variables examined.

3.2. Prediction of the "10% Required" Indicator

As regards the prediction of the "10% required" indicator, the accuracy score ranges from 72.70% to 81.03% (Table 4). Model A, using "house age", "ownership status", and "household size" as input variables, presents a reasonable model, with an accuracy score of 73.05%, good performance metrics, i.e., exceeding 0.73, and a good confusion matrix, i.e., predicting correctly 75% of energy poor and 65% of non-energy poor households. Similarly, satisfactory results are also displayed for Model B (accuracy score of 72.70%).

However, Model C emerges as the best model, as it increases the accuracy score by 8 percentage points compared to Model A, reaching the highest accuracy score (81.03%). Moreover, the performance metrics (Precision, Recall and F-Measure) of Model C are impressively good, exceeding 0.80 for the weighted average. ROC Area is equal to 0.86, almost approaching 1.0, indicating that the examined model is very close to an ideal model. Similarly, according to the diagonal parts of the confusion matrix of Model C, 81% of energy poor households and 80% of non-energy poor households are correctly predicted.

In other words, a neural network using all five (5) input social–geographical variables (house age, ownership status, household size, house area, and elevation) can correctly predict energy poverty—as described by the 10% indicator based on required expenses—at an impressively good level (81.03%).

The neural network of the best model (Model C) included one hidden layer, which consisted of seven nodes. The filter SMOTE was also applied to adjust the relative frequency between minority and majority classes, i.e., double the instances of minority classes, based on the k-nearest neighbor approach (10 nearest neighbors selected). The specific neural network is illustrated in Figure 3.

Table 4. Prediction of "10% required" indicator and confusion matrices (test sets presented).

Input Variables	Precision	Recall	F-Measure	ROC Area	Class	Accuracy	Confusion Matrix	
				Prediction of "10% Required" Indicator				
Model A: House age Ownership status HH size	0.876	0.755	0.811	0.774	Yes	**73.05%**	**75%**	25%
	0.448	0.652	0.531	0.774	No		35%	**65%**
	0.776	0.730	0.745	0.774	(weighted avg)			
Model B: Model A + House area	0.860	0.769	0.812	0.783	Yes	**72.70%**	**77%**	23%
	0.438	0.591	0.503	0.783	No		41%	**59%**
	0.761	0.727	0.740	0.783	(weighted avg)			
Model C: Model B + Elevation	0.929	0.815	0.868	0.856	Yes	**81.03%**	**81%**	19%
	0.568	0.795	0.662	0.856	No		20%	**80%**
	0.844	0.810	0.820	0.856	(weighted avg)			

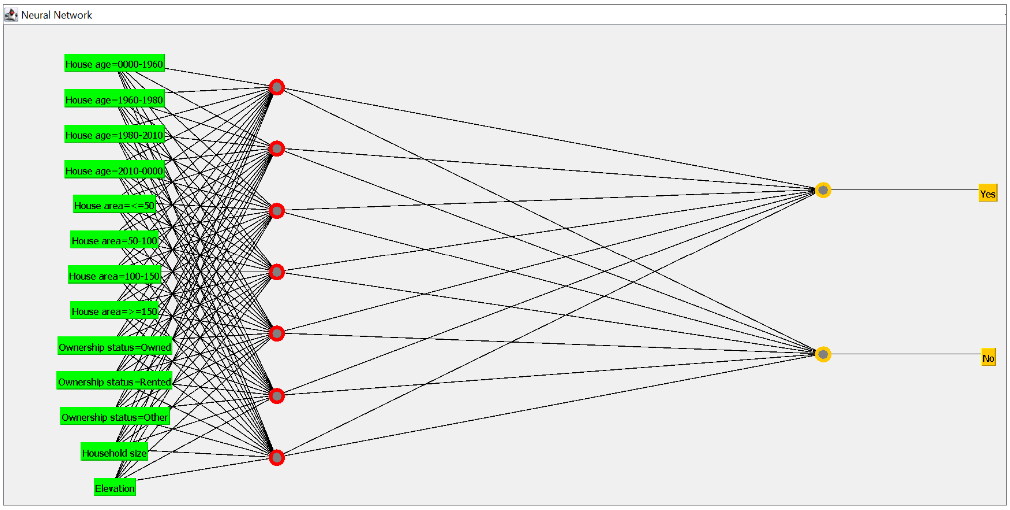

Figure 3. ANN of the model with the best performance, predicting the "10% required" indicator.

3.3. Prediction of the "CEN" Indicator

Regarding the prediction of the "CEN" indicator, all the three models present similar and satisfactory results. Specifically, all of them present good performance metrics (Precision, Recall, F-Measure and ROC Area), i.e., exceeding 0.72 for the weighted average, an accuracy score of 72.91–73.52%, and good scores of the confusion matrix, i.e., predicting accurately 78–82% of energy poor and 64–69% of non-energy poor households, which are considered adequately high. The results of the three models are displayed in Table 5.

Hence, in the case of the "CEN" indicator, a neural network using either only three (3) social variables (house age, ownership status and household size) or four (4) or five (5) variables, i.e., plus house area and elevation, can reliably predict energy poverty– as described by the "CEN" indicator—at a satisfactory level, of the order of 73%.

The neural network of each one of the three models included one hidden layer, which consisted of seven nodes (Figure 4).

Table 5. Prediction of the "CEN" indicator and confusion matrices (test sets presented).

	Prediction of "CEN" Indicator							
Input Variables	**Precision**	**Recall**	**F-Measure**	**ROC Area**	**Class**	**Accuracy**	**Confusion Matrix**	
Model A: House age Ownership status HH size	0.724	0.792	0.757	0.785	Yes	**73.52%**	**79%**	21%
	0.750	0.674	0.710	0.785	No		33%	**67%**
	0.737	0.735	0.734	0.785	(weighted avg)			
Model B: Model A + House area	0.731	0.776	0.753	0.780	Yes	**73.52%**	**78%**	22%
	0.741	0.691	0.715	0.780	No		31%	**69%**
	0.736	0.735	0.735	0.780	(weighted avg)			
Model C: Model B + Elevation	0.707	0.816	0.758	0.783	Yes	**72.91%**	**82%**	18%
	0.761	0.636	0.693	0.783	No		36%	**64%**
	0.733	0.729	0.727	0.783	(weighted avg)			

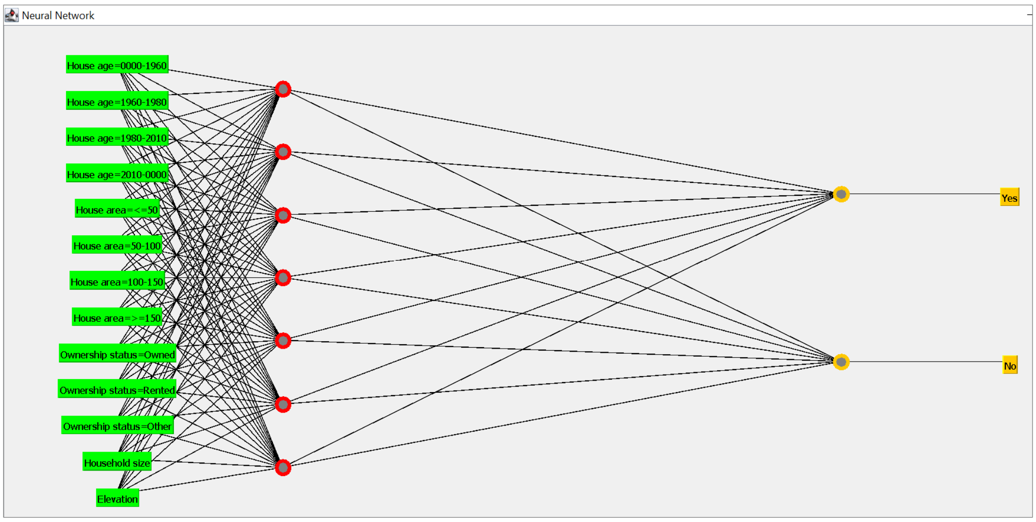

Figure 4. ANN of the model with the best performance, predicting the "CEN" indicator.

3.4. Prediction of the "NEPI" Indicator

As regards the prediction of the "NEPI" indicator, the accuracy score ranges from 71.25% to 82.72% (Table 6). Models A and B present good performance metrics (exceeding 0.7) and high accuracy scores of 71.25–73.94% (Model A and Model B, respectively) but unreasonable rates of the confusion matrix, i.e., only 43–52% of energy poor households are correctly predicted (Model A and Model B, respectively). As a result, on the whole, Models A and B are regarded as non-meaningful models.

On the contrary, Model C presents the best results of the three models, with an accuracy score higher by 9 percentage points compared to Model B (82.72%). The performance metrics (Precision, Recall, F-Measure and ROC Area) of Model C are remarkably good, exceeding 0.82 for the weighted average. ROC Area is equal to 0.88, almost approaching 1.0, indicating that the examined model is very close to an ideal model. Similarly, according to the diagonal parts of the confusion matrix of Model C, 73% of energy poor households and 88% of non-energy poor households are correctly predicted.

In other words, a neural network using all five (5) input social–geographical variables (house age, ownership status, household size, house area, and elevation) can correctly predict energy poverty—as described by the "NEPI" indicator—at a remarkably good level (82.72%).

The neural network of the best model (Model C) included one hidden layer, which consisted of seven nodes. The filter SMOTE was also applied to adjust the relative frequency between minority and majority classes, i.e., double the instances of minority classes, based on the k-nearest neighbor approach (15 nearest neighbors selected). The specific neural network is illustrated in Figure 5.

Table 6. Prediction of the "NEPI" indicator and confusion matrices (test sets presented).

	Prediction of "NEPI" Indicator							
Input Variables	**Precision**	**Recall**	**F-Measure**	**ROC Area**	**Class**	**Accuracy**	**Confusion Matrix**	
Model A: House age Ownership status HH size	0.651	0.433	0.520	0.768	Yes	**71.25%**	**43%**	57%
	0.732	0.869	0.795	0.768	No		13%	**87%**
	0.703	0.712	0.696	0.768	(weighted avg)			
Model B: Model A + House area	0.680	0.520	0.589	0.769	Yes	**73.94%**	**52%**	48%
	0.762	0.863	0.809	0.769	No		14%	**86%**
	0.732	0.739	0.730	0.769	(weighted avg)			
Model C: Model B + Elevation	0.777	0.728	0.752	0.877	Yes	**82.72%**	**73%**	27%
	0.853	0.883	0.867	0.877	No		12%	**88%**
	0.825	0.827	0.826	0.877	(weighted avg)			

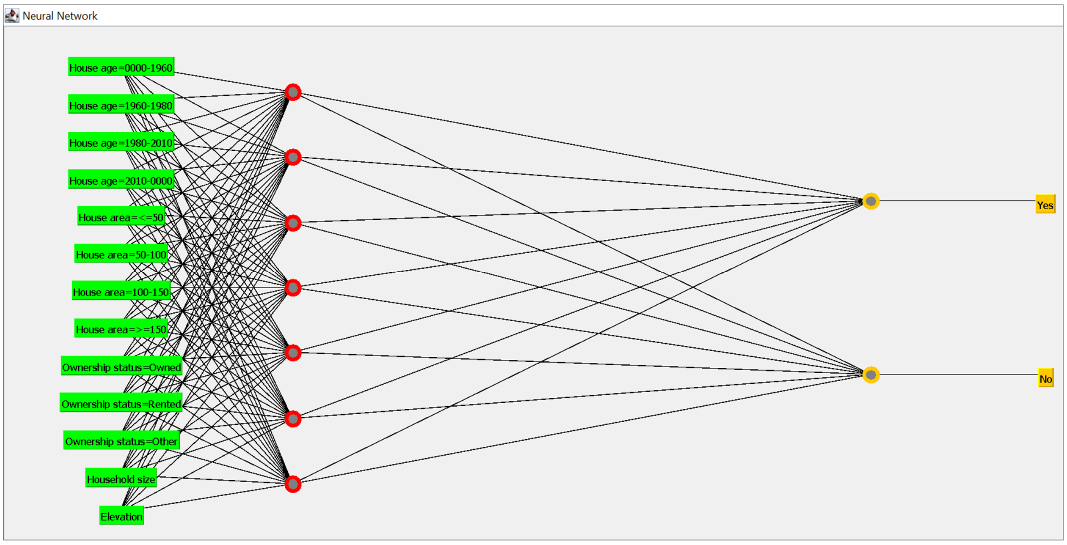

Figure 5. ANN of the model with the best performance, predicting the "NEPI" indicator.

3.5. Prediction of the "IW" Indicator

Regarding the prediction of the "IW" indicator, the accuracy score displays lower rates compared to the previous, objective indicators, varying from 54.19% to 63.98% (Table 7). Models A and B present average performance metrics (marginally exceeding 0.55) and average accuracy scores of 54.19–56.83% (Model B and Model A, respectively) but unreasonable rates of the confusion matrix, i.e., only 43% of energy poor households and only 44% of non-energy poor households are correctly predicted (Model A and Model B, respectively). As a result, on the whole, Models A and B are regarded as non-meaningful models.

On the contrary, Model C presents better results compared to the other two models, with an accuracy score higher by 7 percentage points compared to Model A (63.98%). The performance metrics (Precision, Recall, F-Measure and ROC Area) of Model C are relatively good, exceeding 0.63 for the weighted average. ROC Area is equal to 0.68, implying an acceptable model. Similarly, according to the diagonal parts of the confusion matrix of Model C, 54% of energy poor households and 73% of non-energy poor households are correctly predicted.

In other words, a neural network using all five (5) input social–geographical variables (house age, ownership status, household size, house area, and elevation) can correctly predict energy poverty—as described by the "IW" indicator, expressing the inability to keep a home adequately warm—at an acceptable level (63.98%).

The neural network of the best model (Model C) included one hidden layer, which consisted of seven nodes. The filter SMOTE was also applied to adjust the relative frequency between minority and majority classes, i.e., double the instances of minority classes, based on the k-nearest neighbor approach (15 nearest neighbors selected). The specific neural network is illustrated in Figure 6.

Table 7. Prediction of the "IW" indicator and confusion matrices (test sets presented).

	Prediction of "IW" Indicator							
Input Variables	**Precision**	**Recall**	**F-Measure**	**ROC Area**	**Class**	**Accuracy**	**Confusion Matrix**	
Model A: House age Ownership status HH size	0.549	0.425	0.479	0.603	Yes	**56.83%**	**43%**	57%
	0.579	0.694	0.631	0.603	No		31%	**69%**
	0.565	0.568	0.560	0.603	(weighted avg)			
Model B: Model A + House area	0.508	0.658	0.573	0.562	Yes	**54.19%**	**66%**	34%
	0.594	0.440	0.506	0.562	No		56%	**44%**
	0.554	0.542	0.537	0.562	(weighted avg)			
Model C: Model B + Elevation	0.635	0.538	0.583	0.678	Yes	**63.98%**	**54%**	46%
	0.643	0.729	0.683	0.678	No		27%	**73%**
	0.639	0.640	0.636	0.678	(weighted avg)			

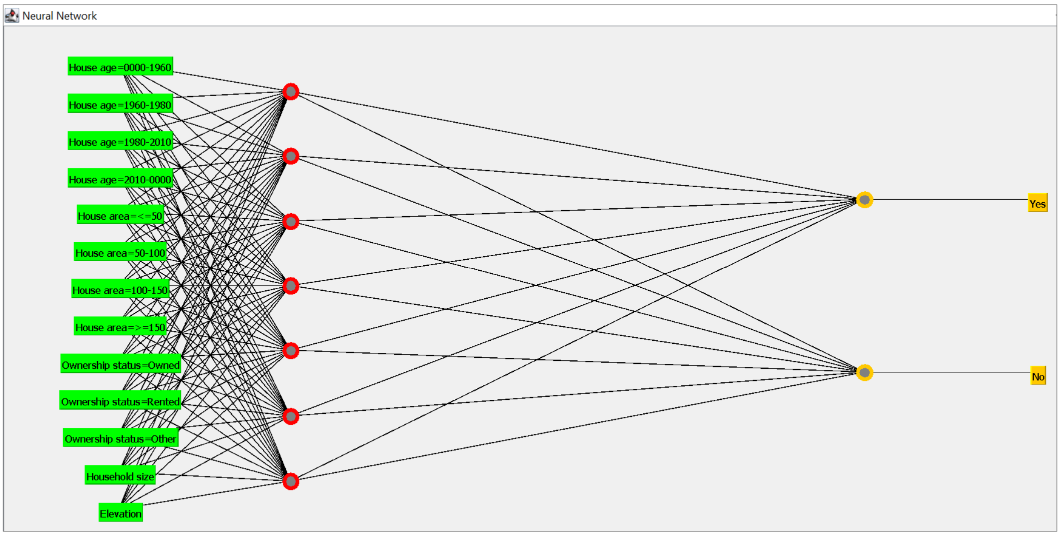

Figure 6. ANN of the model with the best performance, predicting the "IW" indicator.

3.6. Prediction of the "AB" Indicator

As regards the prediction of the "AB" indicator, Model A appears to be inappropriate, especially due to the considerably low values of the confusion matrix, being unable to correctly predict the class of energy poor households (correct prediction of the order of 12%), as shown in Table 8. The other two models present significantly improved results. Model B, with the addition of house area in the input variables, presents an accuracy score higher by 10 percentage points compared to Model A, quite good performance metrics (Precision, Recall, F-Measure and ROC Area), i.e., exceeding 0.66 for the weighted average, as well as satisfactory rates of the confusion matrix (69% of energy poor and 65% of non-energy poor households are correctly predicted).

However, the best results arise in the case of Model C (addition of elevation in input variables), which displays a significant improvement in the performance metrics and also in the confusion matrix elements. Specifically, "F-Measure" (0.754) and "ROC Area" (0.769) exceed 0.75, while also 70% of the energy poor households and 79% of the non-energy poor households are correctly predicted, implying a good and reliable model.

In other words, a neural network using all five (5) input social–geographical variables (house age, ownership status, household size, house area, and elevation) can correctly predict energy poverty—as described by the "AB" indicator, expressing arrears on energy bills—at a satisfactory level (75.35%).

The neural network of the best model (Model C) included one hidden layer, which consisted of seven nodes. The filter SMOTE was also applied to adjust the relative frequency between minority and majority classes, i.e., double the instances of minority classes, based on the k-nearest neighbor approach (15 nearest neighbors selected). The specific neural network is illustrated in Figure 7.

Table 8. Prediction of the "AB" indicator and confusion matrices (test sets presented).

Prediction of "AB" Indicator								
Input Variables	**Precision**	**Recall**	**F-Measure**	**ROC Area**	**Class**	**Accuracy**	**Confusion Matrix**	
Model A: House age Ownership status HH size	0.398	0.122	0.187	0.600	Yes	**57.04%**	**12%**	88%
	0.595	0.875	0.708	0.600	No		13%	**87%**
	0.515	0.570	0.497	0.600	(weighted avg)			
Model B: Model A + House area	0.573	0.693	0.628	0.682	Yes	**66.76%**	**69%**	31%
	0.758	0.650	0.700	0.682	No		35%	**65%**
	0.683	0.668	0.671	0.682	(weighted avg)			
Model C: Model B + Elevation	0.694	0.697	0.696	0,769	Yes	**75.35%**	**70%**	31%
	0.794	0.792	0.793	0.769	No		21%	**79%**
	0.754	0.754	0.754	0.769	(weighted avg)			

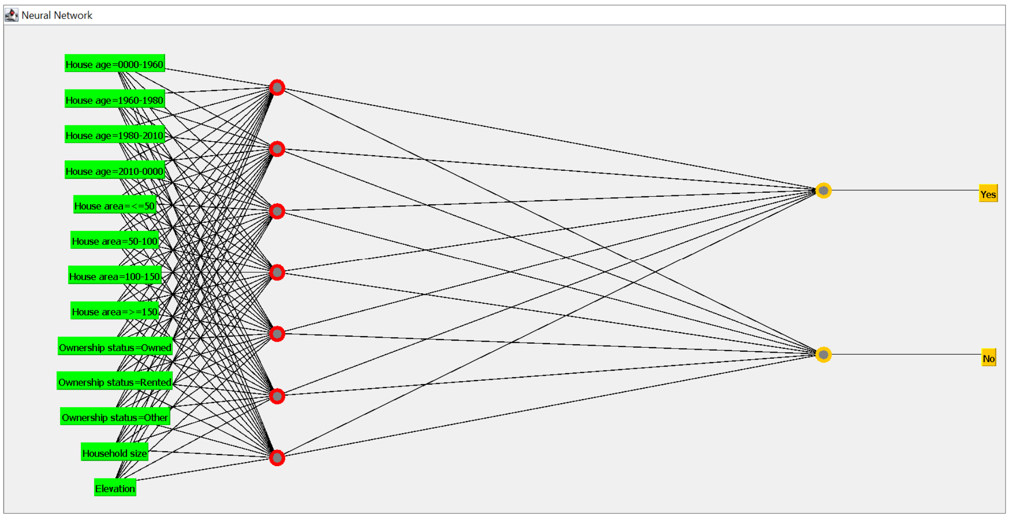

Figure 7. ANN of the model with the best performance, predicting the "AB" indicator.

3.7. Prediction of the "DL" Indicator

Regarding the prediction of the "DL" indicator, Models A and B appear to be inappropriate, especially due to the low values of the confusion matrix, being unable to correctly predict the class of energy poor households (correct prediction of 30% and 27% for Models A and B, respectively), as shown in Table 9.

On the contrary, Model C presents better results compared to the other two models, with an accuracy score higher by 4 percentage points compared to Model B (61.71%). The performance metrics (Precision, Recall, F-Measure and ROC Area) of Model C are relatively good, exceeding 0.61 for the weighted average. ROC Area is equal to 0.65, implying an acceptable model. Similarly, according to the diagonal parts of the confusion matrix of

Model C, 64% of energy poor households and 60% of non-energy poor households are correctly predicted.

In other words, a neural network using all five (5) input social–geographical variables (house age, ownership status, household size, house area, and elevation) can correctly predict energy poverty—as described by the "DL" indicator, expressing problems of a leaking roof, damp walls/floors/foundation, or rot in window frames or floor—at an acceptable level (61.71%).

The neural network of the best model (Model C) included one hidden layer, which consisted of seven nodes. The filter SMOTE was also applied to adjust the relative frequency between minority and majority classes, i.e., increasing by 50% the instances of minority classes, based on the k-nearest neighbor approach (15 nearest neighbors selected). The specific neural network is illustrated in Figure 8.

Table 9. Prediction of the "DL" indicator and confusion matrices (test sets presented).

Prediction of "DL" Indicator								
Input variables	**Precision**	**Recall**	**F-Measure**	**ROC Area**	**Class**	**Accuracy**	**Confusion Matrix**	
Model A: House age Ownership status HH size	0.513	0.302	0.380	0.543	Yes	**56.12%**	**30%**	70%
	0.578	0.770	0.660	0.543	No		23%	**77%**
	0.549	0.561	0.535	0.543	(weighted avg)			
Model B: Model A + House area	0.547	0.275	0.366	0.589	Yes	**57.52%**	**27%**	73%
	0.583	0.817	0.681	0.589	No		18%	**82%**
	0.567	0.575	0.540	0.589	(weighted avg)			
Model C: Model B + Elevation	0.562	0.639	0.598	0.648	Yes	**61.71%**	**64%**	36%
	0.674	0.599	0.634	0.648	No		40%	**60%**
	0.624	0.617	0.618	0.648	(weighted avg)			

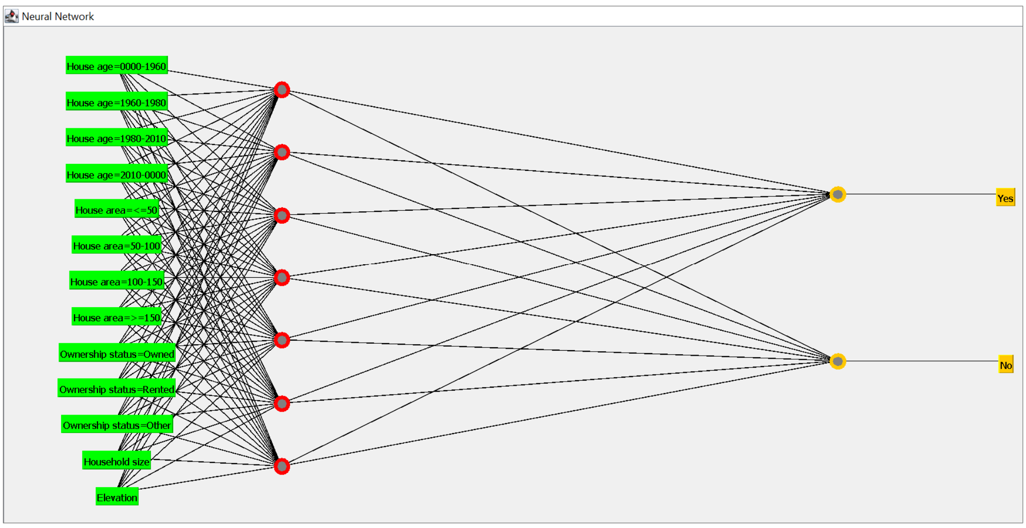

Figure 8. ANN of the model with the best performance, predicting the "DL" indicator.

4. Conclusions

The present paper explores the ability of Artificial Intelligence and, particularly, of Artificial Neural Networks (ANNs) to successfully predict energy poverty after proper "training", a venture endeavored for the first time for the energy poverty issue in Greece, to the best of authors' knowledge. The analysis included the prediction of seven (7) energy poverty indicators (output indicators) based on certain socio-economic/geographical factors (input variables), via training an Artificial Neural Network (ANN), i.e., the Multilayer Perceptron. Three (3) basic models (Model A, Model B and Model C) were tested in terms of combinations of the input variables in order to examine valuable models.

The analysis showed that ANNs managed to predict energy poverty at a remarkably good level of accuracy. All in all, Model C appeared to be the best one among the three (3) models examined for all indicators ("10% actual" indicator, "10% required" indicator, "CEN" indicator, "NEPI" indicator, "IW" indicator, "AB" indicator and "DL" indicator). Specifically, it was shown that an ANN using all five (5) input social–geographical variables (house age, ownership status, household size, house area, and elevation) can correctly predict energy poverty as described by each one of the examined indicators, at a highly satisfactory level, ranging from 61.71% (lowest value for the "DL" indicator) up to 82.72% (highest value for the "NEPI" indicator). Actually, the ANNs of two of the examined indicators, i.e., the "10% required" indicator and the "NEPI" indicator, proved to be impressively accurate, managing to predict energy poverty with a success rate of over 80%, which is considered particularly high in terms of social indicators.

Moreover, it should be noted that, although hundreds of different types of ANNs were tested, the simplest ones were those that finally gave the best results, e.g., by selecting one hidden layer instead of two or more with a large number of nodes each. Moreover, the filter "SMOTE" needed to be applied to almost all cases except for the "CEN" indicator (always via numerous tests of the filter's elements) in order to adjust the relative frequency between minority and majority classes and, hence, mitigate bias in the results.

These findings shed light on an innovative approach of studying and analyzing energy poverty, that of machine learning and, particularly, of ANNs. The high accuracy scores that came up on the examined cases confirmed that ANNs can be a useful tool towards a deeper understanding of the energy poverty roots, which in turn can help decision-making and lead to more targeted policy measures. For instance, the present research showed that if five (5) particular social–geographical variables, i.e., house age, ownership status, household size, house area and elevation, are given for a household, it can be easily estimated whether the household is at risk of some kind of energy poverty. In this way, ANNs can be a useful tool in the hands of the Greek government to identify groups of people more vulnerable to energy poverty, so that they can receive enhanced social support, e.g., allowances, tax reliefs, motivations to renovate their residence, lower energy prices, and so on.

Future research could possibly expand the research of ANNs to further energy poverty variables and indicators or explore more machine learning techniques, other than ANNs, in the same case study. In any case, Artificial Intelligence seems to be a promising tool in understanding the complex nature and the various dimensions of energy poverty.

Author Contributions: Conceptualization, D.K. and L.P.; methodology, L.P. and D.K.; formal analysis, L.P.; resources, L.P.; data curation, L.P.; writing—original draft preparation, L.P. and D.K.; writing—review and editing, L.P.; supervision, D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. IEA. SDG7: Data and Projections. IEA, 2023. Available online: https://www.iea.org/reports/sdg7-data-and-projections (accessed on 20 April 2024).
2. Papada, L.; Kaliampakos, D. A Stochastic Model for energy poverty analysis. *Energy Policy* **2018**, *116*, 153–164. [CrossRef]
3. Eurostat. People at Risk of Poverty or Social Exclusion. 2023. Available online: https://ec.europa.eu/eurostat/databrowser/view/sdg_01_10/default/table?lang=en (accessed on 25 April 2024).
4. Kalfountzou, E.; Tourkolias, C.; Mirasgedis, S.; Damigos, D. Identifying Energy-Poor Households with Publicly Available Information: Promising Practices and Lessons Learned from the Athens Urban Area, Greece. *Energies* **2024**, *17*, 919. [CrossRef]
5. Atsalis, A.; Mirasgedis, S.; Tourkolias, C.; Diakoulaki, D. Fuel poverty in Greece: Quantitative analysis and implications for policy. *Energy Build.* **2016**, *131*, 87–98. [CrossRef]
6. Papada, L.; Kaliampakos, D. Measuring energy poverty in Greece. *Energy Policy* **2016**, *94*, 157–165. [CrossRef]

7. Papada, L.; Kaliampakos, D. Energy poverty in Greek mountainous areas: A comparative study. *J. Mt. Sci.* **2017**, *14*, 1229–1240. [CrossRef]
8. Palmos Analysis. Thessaloniki: 190,000 Households Vulnerable or in a State of Energy Poverty. Available online: https://parallaximag.gr/life/energiaki-ftochia-stopoleodomiko-sigkrotima-thessalonikis (accessed on 20 April 2024).
9. Boemi, S.-N.; Avdimiotis, S.; Papadopoulos, A.M. Domestic energy deprivation in Greece: A field study. *Energy Build.* **2017**, *144*, 167–174. [CrossRef]
10. Ntaintasis, E.; Mirasgedis, S.; Tourkolias, C. Comparing different methodological approaches for measuring energy poverty: Evidence from a survey in the region of Attika, Greece. *Energy Policy* **2019**, *125*, 160–169. [CrossRef]
11. Spiliotis, E.; Arsenopoulos, A.; Kanellou, E.; Psarras, J.; Kontogiorgos, P. A multi-sourced data based framework for assisting utilities identify energy poor households: A case-study in Greece. *Energy Sources Part B Econ. Plan. Policy* **2020**, *15*, 49–71. [CrossRef]
12. Papada, L.; Kaliampakos, D. Being forced to skimp on energy needs: A new look at energy poverty in Greece. *Energy Res. Soc. Sci.* **2020**, *64*, 101450. [CrossRef]
13. Lyra, K.; Mirasgedis, S.; Tourkolias, C. From measuring fuel poverty to identification of fuel poor households: A case study in Greece. *Energy Effic.* **2022**, *15*, 6. [CrossRef]
14. Kalfountzou, E.; Papada, L.; Damigos, D.; Degiannakis, S. Predicting energy poverty in Greece through statistical data analysis. *Int. J. Sustain. Energy* **2022**, *41*, 1605–1622. [CrossRef]
15. Halkos, G.; Kostakis, I. Exploring the persistence and transience of energy poverty: Evidence from a Greek household survey. *Energy Effic.* **2023**, *16*, 50. [CrossRef]
16. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2019.
17. Hong, Z.; Park, I.K. Comparative Analysis of Energy Poverty Prediction Models Using Machine Learning Algorithms. *JKPA* **2021**, *56*, 239–255. [CrossRef]
18. Pino-Mejías, R.; Pérez-Fargallo, A.; Rubio-Bellido, C.; Pulido-Arcas, J.A. Artificial neural networks and linear regression prediction models for social housing allocation: Fuel Poverty Potential Risk Index. *Energy* **2018**, *164*, 627–641. [CrossRef]
19. Hassani, H.; Yeganegi, M.R.; Beneki, C.; Unger, S.; Moradghaffari, M. Big Data and Energy Poverty Alleviation. *BDCC* **2019**, *3*, 50. [CrossRef]
20. Benardos, A. Artificial intelligence in underground development: A study of TBM performance. In Proceedings of the UNDERGROUND SPACES 2008, New Forest, UK, 8–10 September 2008; pp. 21–32. [CrossRef]
21. Moon, J.W.; Kim, J.-J. ANN-based thermal control models for residential buildings. *Build. Environ.* **2010**, *45*, 1612–1625. [CrossRef]
22. Seo, J.; Kim, S.; Lee, S.; Jeong, H.; Kim, T.; Kim, J. Data-driven approach to predicting the energy performance of residential buildings using minimal input data. *Build. Environ.* **2022**, *214*, 108911. [CrossRef]
23. Michailidis, P.; Michailidis, I.T.; Gkelios, S.; Karatzinis, G.; Kosmatopoulos, E.B. Neuro-distributed cognitive adaptive optimization for training neural networks in a parallel and asynchronous manner. *ICA* **2023**, *31*, 19–41. [CrossRef]
24. Elbeltagi, E.; Wefki, H. Predicting energy consumption for residential buildings using ANN through parametric modeling. *Energy Rep.* **2021**, *7*, 2534–2545. [CrossRef]
25. Tu, J.V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J. Clin. Epidemiol.* **1996**, *49*, 1225–1231. [CrossRef]
26. Papada, L.; Kaliampakos, D. Exploring Energy Poverty Indicators Through Artificial Neural Networks. In *Artificial Intelligence and Sustainable Computing*; Pandit, M., Gaur, M.K., Rana, P.S., Tiwari, A., Eds.; Algorithms for Intelligent Systems; Springer Nature: Singapore, 2022; pp. 231–242. [CrossRef]
27. Fausett, L.V. *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*; Pearson Education: New Delhi, India, 2008.
28. Sietsma, J.; Dow, R.J.F. Creating artificial neural networks that generalize. *Neural Netw.* **1991**, *4*, 67–79. [CrossRef]
29. Rajić, M.N.; Milovanović, M.B.; Antić, D.S.; Maksimović, R.M.; Milosavljević, P.M.; Pavlović, D.L. Analyzing energy poverty using intelligent approach. *Energy Environ.* **2020**, *31*, 1448–1472. [CrossRef]
30. Cheng-wen, Y.; Jian, Y. Application of ANN for the prediction of building energy consumption at different climate zones with HDD and CDD. In *2010 2nd International Conference on Future Computer and Communication*; IEEE: Wuhan, China, 2010; pp. V3-286–V3-289. [CrossRef]
31. Pavićević, M.; Popović, T. Forecasting Day-Ahead Electricity Metrics with Artificial Neural Networks. *Sensors* **2022**, *22*, 1051. [CrossRef]
32. Heghedus, C.; Segarra, S.; Chakravorty, A.; Rong, C. Neural Network Architectures for Electricity Consumption Forecasting. In *2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*; IEEE: Atlanta, GA, USA, 2019; pp. 776–783. [CrossRef]
33. Tardioli, G.; Kerrigan, R.; Oates, M.; O'Donnell, J.; Finn, D. A Data-Driven Modelling Approach for Large Scale Demand Profiling of Residential Buildings. In Proceedings of the BS 2017: Conference of International Building Performance Simulation Association, San Francisco, CA, USA, 7–9 August 2017; Barnaby, C.S., Wetter, M., Eds.; International Building Performance Simulation Association: Rapid City, SD, USA, 2017.

34. López-Vargas, A.; Ledezma-Espino, A.; Sanchis-de-Miguel, A. Methods, data sources and applications of the Artificial Intelligence in the Energy Poverty context: A review. *Energy Build.* **2022**, *268*, 112233. [CrossRef]
35. Longa, F.D.; Sweerts, B.; Van Der Zwaan, B. Exploring the complex origins of energy poverty in The Netherlands with machine learning. *Energy Policy* **2021**, *156*, 112373. [CrossRef]
36. Al Kez, D.; Foley, A.; Abdul, Z.K.; Del Rio, D.F. Energy poverty prediction in the United Kingdom: A machine learning approach. *Energy Policy* **2024**, *184*, 113909. [CrossRef]
37. Bienvenido-Huertas, D.; Pérez-Fargallo, A.; Alvarado-Amador, R.; Rubio-Bellido, C. Influence of climate on the creation of multilayer perceptrons to analyse the risk of fuel poverty. *Energy Build.* **2019**, *198*, 38–60. [CrossRef]
38. Bienvenido-Huertas, D.; Sánchez-García, D.; Marín-García, D.; Rubio-Bellido, C. Analysing energy poverty in warm climate zones in Spain through artificial intelligence. *J. Build. Eng.* **2023**, *68*, 106116. [CrossRef]
39. Abbas, K.; Butt, K.M.; Xu, D.; Ali, M.; Baz, K.; Kharl, S.H.; Ahmed, M. Measurements and determinants of extreme multidimensional energy poverty using machine learning. *Energy* **2022**, *251*, 123977. [CrossRef]
40. Balaskas, A.; Papada, L.; Katsoulakos, N.; Damigos, D.; Kaliampakos, D. Energy poverty in the mountainous town of Metsovo, Greece. *J. Mt. Sci.* **2021**, *18*, 2240–2254. [CrossRef]
41. Papada, L.; Balaskas, A.; Katsoulakos, N.; Kaliampakos, D.; Damigos, D. Fighting Energy Poverty Using User-Driven Approaches in Mountainous Greece: Lessons Learnt from a Living Lab. *Energies* **2021**, *14*, 1525. [CrossRef]
42. Karani, I.; Papada, L.; Kaliampakos, D. Energy poverty signs in mountainous Greek areas: The case of Agrafa. *Int. J. Sustain. Energy* **2022**, *41*, 1408–1433. [CrossRef]
43. Roberts, D.; Vera-Toscano, E.; Phimister, E. Fuel poverty in the UK: Is there a difference between rural and urban areas? *Energy Policy* **2015**, *87*, 216–223. [CrossRef]
44. DECC (Dept for Energy and Climate Change). Annual Fuel Poverty Statistics Report 2015. DECC: London, UK, 2015. Available online: https://assets.publishing.service.gov.uk/media/5a814b3340f0b6230269684c/Fuel_Poverty_Report_2015.pdf (accessed on 25 April 2024).
45. Annual Progress Report of the Action Plan for Combating Energy Poverty Year 2021 (in Greek). Ministry of Environment & Energy, 2021. Available online: https://ypen.gov.gr/wp-content/uploads/2022/04/SDEE-Annual-report-2021-v4-14032022-clean.pdf (accessed on 15 April 2024).
46. EU Statistics on Income and Living Conditions. Eurostat, 2024. Available online: https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions (accessed on 25 April 2024).
47. 'Arrears on Utility Bills'. Eurostat, 2024. Available online: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20200120-1 (accessed on 25 April 2024).
48. 'EU Statistics on Income and Living Conditions (EU-SILC) Methodology—Housing Deprivation'. Eurostat, 2024. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_housing_deprivation (accessed on 25 April 2024).
49. European Commission. *An Energy Policy for Customers. Commission Staff Working Paper*; European Commission: Brussels, Belgium, 2010.
50. Papada, L.; Kaliampakos, D. Developing the energy profile of mountainous areas. *Energy* **2016**, *107*, 205–214. [CrossRef]
51. Frank, E.; Hall, M.A.; Witten, I.H. *The WEKA Workbench. Data Mining. Practical Machine Learning Tools and Techniques*, 4th ed.; Morgan Kaufmann: Waikato, New Zealand, 2016.