Ca' Foscari University of Venice

Department of Economics

Working Paper

Antonella Basso
Marco Corazza
Lorenzo Tonon

Recurrent Neural Networks for real estate evaluation in the Italian market

ISSN: 1827-3580
No. 22/WP/2026

Working Papers
Department of Economics
Ca' Foscari University of Venice
No. 22/WP/2026
ISSN 1827-3580

Recurrent Neural Networks for real estate evaluation in the Italian market

Antonella Basso
Ca' Foscari University of Venice

Marco Corazza
Ca' Foscari University of Venice

Lorenzo Tonon
Ca' Foscari University of Venice

Abstract
Accurate evaluation of real estate prices represents a complex task, that holds significant relevance for the majority of the actors involved in a society, as homebuyers/sellers, investors, and policymakers. Over the last decades, in parallel with technological advancements, scholars and practitioners proposed several models and methodologies, spanning from simple hedonic regression models to sophisticated machine learning algorithms, in the attempt to improve the accuracy of price forecasts. Nonetheless, the interest towards the development of more precise and performing methodologies remains high, not only for static valuation but also, more recently, for dynamic approaches. Focusing on dynamic valuation, we turn the attention towards the Recurrent Neural Network (RNN) models: while widely tested for stock prices forecast, RNNs have seen limited to no application in the domain of real estate valuation. In this study, using official data on prices of Italian residential properties, we propose and implement two different architectures for RNN models to perform both price prediction and relevance analysis with respect to the explanatory variables. Differently from other studies, we utilize data for the whole Italian territory, focusing, rather than on individual properties' valuation, on a range of prices for groups of properties identified by house type, geographical position and state of conservation. Specifically, the output variables of our analysis are represented by the minimum and maximum prices that determine this range. The results of our analysis indicate that a Standard Recurrent Neural Network forecasts well the prices in small municipalities (with less than 15 000 inhabitants), while a Double-Input Layer RNN, with an input layer dedicated to static features and a separate input layer for dynamic features, performs better in medium-sized (with a number of inhabitants between 15 000 and 50 000 inhabitants) and large municipalities (with more than 50 000 municipalities).

Keywords
Real Estate valuation, House price forecast, Recurrent Neural Networks, Machine Learning, LSTM Networks

JEL Codes
C45, G17, R31, R32

Address for correspondence:
Antonella Basso
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
e-mail: basso@unive.it

The Working Paper Series
is available only on line
(https://www.unive.it/dec-workingpapers)
For editorial correspondence, please contact:
wp.dse@unive.it

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta San Giobbe
30121 Venice Italy
Fax: ++39 041 2349210

Recurrent Neural Networks for real estate evaluation in the Italian market

Antonella Basso, Marco Corazza, Lorenzo Tonon

[1]Department of Economics, Ca' Foscari University of Venice, Cannaregio 873, Fondamenta San Giobbe, Venice, 30121, VE, Italy.

Abstract

Accurate evaluation of real estate prices represents a complex task, that holds significant relevance for the majority of the actors involved in a society, as home-buyers/sellers, investors, and policymakers. Over the last decades, in parallel with technological advancements, scholars and practitioners proposed several models and methodologies, spanning from simple hedonic regression models to sophisticated machine learning algorithms, in the attempt to improve the accuracy of price forecasts. Nonetheless, the interest towards the development of more precise and performing methodologies remains high, not only for static valuation but also, more recently, for dynamic approaches.

Focusing on dynamic valuation, we turn the attention towards the Recurrent Neural Network (RNN) models: while widely tested for stock prices forecast, RNNs have seen limited to no application in the domain of real estate valuation. In this study, using official data on prices of Italian residential properties, we propose and implement two different architectures for RNN models to perform both price prediction and relevance analysis with respect to the explanatory variables.

Differently from other studies, we utilize data for the whole Italian territory, focusing, rather than on individual properties' valuation, on a range of prices for groups of properties identified by house type, geographical position and state of conservation. Specifically, the output variables of our analysis are represented by the minimum and maximum prices that determine this range. The results of our analysis indicate that a Standard Recurrent Neural Network forecasts well the prices in small municipalities (with less than 15 000 inhabitants), while a Double-Input Layer RNN, with an input layer dedicated to static features and a separate input layer for dynamic features, performs better in medium-sized (with a number of inhabitants between 15 000 and 50 000 inhabitants) and large municipalities (with more than 50 000 municipalities).

Keywords: Real Estate valuation, House price forecast, Recurrent Neural Networks, Machine Learning, LSTM Networks

JEL Classification: C45 , G17 , R31 , R32

1 Introduction

A correct understanding and evaluation of real estate prices constitutes a fundamental factor both for individual investment decisions and overall economic stability. The housing market represents a non-irrelevant component of the overall economy, and its impact reaches beyond private household wealth, to public policy choices and economic growth in general. An accurate prediction and analysis of housing prices could, therefore, have significant consequences for anyone participating directly or indirectly to the market, be it policymakers making regulatory choices related to housing or price control, financial institutions involved in managing investment portfolios and assessing risks, or other participants. In a similar fashion, accurate housing prices forecast would significantly facilitate the decision-making processes of home-buyers, investors and developers [1, 2].

Over the last decades, the Italian housing market has been experiencing noticeable fluctuations due to a mix of local and global factors linked, for example, to economic recession, demographic shifts, and rapidly changing governmental regulations. When trying to analyse such movements, one of the major concerns is related to the fact that real estate markets within Italy, especially at the municipal level, are disparate in their nature. Italy is, indeed, a country with great regional economic differences: each municipality has its peculiar mix of socio-cultural-economic traits, such as employment situation, available services, and geographical characteristics. In this respect, this heterogeneity among local markets calls into question the task of building a single house price forecast model. A valuation model performing well in the city center of Milan can fail when tested on countryside municipalities in southern Italy.

Until now, traditional econometric approaches to real estate valuation have mainly adopted solutions as the hedonic pricing model [3] and the Auto-Regressive Integrated Moving-Average (ARIMA) methods [4, 5], in order to estimate house prices on the basis of various factors such as location, size, and other economic variables [6, 7]. However, these models are not capable of capturing the complex non-linear relationships and high-dimensional time dependencies that often characterize the housing market. Indeed, the structural rigidity of traditional models prevents them from taking complete advantage of the temporal patterns shown by data, making it difficult for them to cope with the rapid changes in market dynamics, especially when dealing with geographically disperse and diverse markets, like Italy.

With the advent of sophisticated machine learning techniques, several different opportunities started to arise for enhanced predictive accuracy in different financial sectors [8], including real estate [9]. Amongst these techniques, static artificial neural networks such as feedforward neural networks (FNNs) and multi-layer perceptrons (MLPs) have started to gain interest due to their ability in modeling complex non-linear relationships between variables [10]. On the other hand, though capable of modelling non-linear interactions, static neural networks are limited by their structural inability to utilize and exploit sequential data or temporal dependencies, which represent a fundamental aspect while performing time series data analysis and housing prices forecast. Indeed, although static neural networks have been widely proven to be capable of making reasonably accurate point estimates of house prices by exploiting the most recent market data [11–14], they lack predictive ability for long-term future price changes as they are not intrinsically designed to take into consideration the time-series nature of the data. Housing prices are, anyway, intrinsically driven by historical events like trends, discontinuities in the economy, and changes in policy that occur over time [15, 16]. Disregarding the chronological order in which these events occur could potentially lead to poor long-term forecasting, especially in fields like real estate, where past prices and historical fluctuations are often regarded as crucial indicators for such predictions.

Recurrent Neural Networks (RNNs) were developed and proposed exactly to address this limitation, both in the form of Long Short-term Memory (LSTM) networks [17] and Bidirectional Recurrent Neural (BRNN) networks [18]. RNNs are specifically designed to model sequential and temporal data [17] in an efficient and structured way. The main difference between RNNs and static neural networks lies in the fact that the former have an internal state that keeps some sort of "memory" regarding previous inputs, allowing them to effectively capture temporal dependencies. Because of this property, and given what we discussed about the house prices nature, RNNs appear as potentially ideal candidates for house price prediction, as they should be able to learn from the short-term volatility of prices and long-term trends, incorporating past values into their predictions [19].

Anyway, differently from static neural networks, Recurrent Neural Networks have seen limited application and utilization in financial and economic studies, with the majority of those focusing on stock prices and volatility prediction [19–21].For example, [21] proposed an attention-based LSTM model for financial time series forecasting, that showed a significantly enhanced predictive performance thanks to its ability to focus on the most relevant information in the input sequence.

Despite the strong results in stock and commodity price prediction, RNNs applications in real estate valuation are relatively scarce. A limited number of studies [9, 22–24], have explored the use of LSTM models to forecast real estate prices. [9], for instance, demonstrated how LSTM networks could effectively capture market trends and adequately forecast residential real estate selling prices with reasonable accuracy, especially when compared to traditional regression models. [24] explored the utilization of RNNs and boosting algorithms for property appraisal, discussing that, while RNNs improved the accuracy of predictions, further research is necessary, with a specific attention on the adaptation of these models to the complexities of real estate markets, such as regional variations and macroeconomic influences. However, the literature addressing real estate valuation through RNNs remains limited, and challenges related, for example, to the usage of location-based features, neighborhood effects, or data scarcity are often overlooked [24].

One of the key advantages of RNNs lies in their ability to model non-linear relationships between variables and, as said, capture temporal patterns. RNN models, nonetheless, still present some key limitations and drawbacks, mainly linked to the large amount of high-quality data needed for optimal training. Especially in niche financial markets and in real estate markets such data are often insufficient or unavailable, with this resulting in over-fitting issues and poor training [16]. In addition to this, the interpretability of the results and outputs of RNNs remains a difficult task, making their efficacy more limited for fields where transparency and explicability of results constitutes a fundamental component of the analysis. Despite their limitations, anyway, RNNs still represent a viable and interesting opportunity for research in various domains, including financial valuation.

The research question of the present study is centered around the analysis and application of Recurrent Neural Networks to real estate price prediction in different Italian municipalities: given the pronounced internal regional differences that characterize the Italian territory, this task represents an exemplary test case to assess the capability of RNNs to forecast real estate prices.

Most importantly, the integration of RNNs into real estate valuation, not only expands the methodological toolkit of economists and data scientists, but also provides practical benefits from the perspective of market participants. In this way, on the one hand, investors would, indeed, become able to anticipate market movements better, reducing risks and improving portfolio management strategies. On the other hand, policymakers might be able to use such projections to assess the impacts of economic policy choices on affordability of residential properties so as to develop intervention policies aimed at reducing housing market inequities between different regions or citizens. For individual home-buyers, accurate forecasts would help them make informed and well-reasoned decisions on market entry or market exit and thus optimize household financial planning.

Besides the methodological novelty, the employment of RNNs presented in this paper aims to contribute to the wider debate on how machine learning can be used in regional economics and financial valuation. The granularity of the Italian housing market data with differentiation at the municipality level creates a unique environment to test the ability of RNNs for highly localized contexts. Moreover, the Italian market presents an additional advantage linked to the availability and transparency of real estate data provided by the Osservatorio del Mercato Immobiliare (OMI). The OMI database, managed by the Agenzia delle Entrate (AdE), includes extensive and updated data relating to sales and valuations of properties across all municipalities in Italy. This huge information source provides official standardized quotations of prices and market data that can be reliably used in econometric and financial analysis, drastically reducing the information asymmetry and greatly improving the transparency within the market.

The present study aims also at explaining the complex relationships among specific regional- and municipal-level economic indicators and residential property values, in order to improve the overall understanding of the dynamics intrinsic to the (Italian) housing market. Furthermore, we present a replicable and flexible model that could well be tested and used for prediction and analysis in other regions or nations characterized by similar challenges.

In conclusion, the current study investigates the efficiency of Recurrent Neural Networks in modelling house prices for Italian municipalities by providing, on empirical grounds, an approach that overcomes drawbacks and gives real estate value estimation in a diverse and multifaceted market. This study attempts to understand whether RNNs can outperform traditional predictive models.

The remainder of this paper is organized as follows. Section 2 introduces the methodology and presents the RNN models proposed, namely a straightforward RNN model and a double-headed RNN model. Section 3 discusses the data used, their transformation and pre-processing, and the procedures used for data cleaning and outlier detection. Section 4 presents and discusses the results of the empirical investigation of minimum and maximum house prices carried out on Italian municipalities; moreover, an analysis of the input relevance is also undertaken to allow for a better economic and financial interpretation of the results. Finally, Section 5 proposes some concluding remarks.

2 Methodology

2.1 Theoretical and Mathematical Considerations

Real estate valuation is often regraded as a complex task because of the multifaceted nature of the factors that influence the price of properties. Historically, many econometric estimation techniques have been widely applied for forecasting residential property values: among them, hedonic pricing models and ARIMA models. These models fail to accurately illustrate the intricate, nonlinear relationships among the variables and usually neglect the temporal dependence that defines time series data, such as house prices. In light of the inefficiency of the previously discussed approaches, ANNs, or more specifically RNNs, offer a very serious alternative, as they can learn complex nonlinear patterns with ease and handle time-dependent data with great success.

Neural Networks: Structure and Suitability

An Artificial Neural Network (ANN) is a computational model designed to replicate and simulate the manner in which the human brain processes and analyses data. In a standard feedforward neural network (FNN), the input features, which in our context span from property characteristics to economic indicators and geographic information, are processed and passed through several layers of computational nodes, called neurons. Each layer transforms the input exploiting a dynamic set of weights and biases, producing a high number of intermediate outputs until the final house price is predicted. This transformation process is governed by a simple set of equation of the following form, for each neuron in a given layer l:

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)}$$

where $W^{(l)}$ is a matrix containing a list of weights for the l-th layer, $a^{(l-1)}$ is the activation of the previous layer, and $b^{(l)}$ is the bias vector. The output of each layer is then passed through a non-linear activation function $f(\cdot)$, which is used to model non-linear relationships between variables:

$$a^{(l)} = f(z^{(l)})$$

The most widely used functions include the rectified linear unit (ReLU), the sigmoid, and the hyperbolic tangent (tanh) activation functions, and enable the network to learn from complex interactions between inputs and outputs. This represents a key factor in the context of house price prediction, given that it allows the model to identify non-linear relationships between variables such as interest rates, property characteristics, and regional economic factors.

Among the various architectures of neural networks, Multi-Layer Perceptron is one of the most common structures used for feedforward neural networks. An MLP includes multiple layers of artificial neurons, organized in an input layer, one or more hidden layers, and a final output layer. In housing price prediction, input is a set of features that would include property features, economic measures, and geographic data, whereas the output is the predicted house price or value.

MLPs and, in general, Feed-forward neural networks, while successful in non-linear interactions modeling, show some limitations when applied to the analysis of time series data. Housing prices are intrinsically influenced by historical values and tendencies; hence, temporal dependencies present an intrinsic feature of any robust predictive model. However, static neural networks consider data independently of its order, and thus it brings insufficient capability in handling time-dependent information. This constraint hence requires the usage of more advanced architectures that are capable of learning from sequences of data efficiently, which include RNNs.

Recurrent Neural Networks: Modeling Temporal Dependencies

Recurrent Neural Networks are explicitly designed to process sequential information and are particularly suitable for tasks related to time series forecasting, such as dynamic house price prediction. Unlike feedforward networks, that allow data to flow only in one way, RNNs have several loops in their architecture, hence enabling the model to maintain the internal state or memory which captures information from past time steps. This memory allows the RNNs to summarize past information in order to make predictions coherent with the observed temporal dynamics and patterns, making them appropriate for applications where outcomes depend on historical patterns.

The core of a RNN is the hidden state h_t, which represents the memory of the network at a certain time step t. At each subsequent time step, the hidden state is updated based on the current input x_t and the previous hidden state h_{t-1}:

$$h_t = f(W_h h_{t-1} + W_x x_t + b)$$

where W_h is the matrix containing the weights used in the previous hidden state, W_x is the weight matrix for the current input, and b is the bias. The non-linear activation function $f(\cdot)$ is used to make sure that the model is able to learn the non-linear dependencies between past and current data.

As we can see, this recursion-based structure allows RNNs to memorize and process sequential information, which makes them perfectly suitable for our task, where temporal dependencies are crucial. For example, the house price in a specific region might depend on historical sales, regional economic growth, and other wider market conditions. Exploiting the memory about previous inputs, RNNs can capture these dependencies better and yield more accurate forecasts. Still, RNNs, show several noticeable drawbacks, which complicate the learning process of these long-term dependencies. When a network is run through several time steps, the gradients at training become smaller and smaller, thereby making it difficult for the model to effectively perform the weight update. This is a well-known problem called the "vanishing gradient problem"; it partially restricts RNNs from capturing information across long sequences of data, a key aspect in house price prediction, where long-run trends can become permanent and show their effects even after a large number of years.

Long Short-Term Memory Networks (LSTMs): Addressing Long-Term Dependencies

Long Short-Term Memory (LSTM) networks, a common variant of RNNs, have been studied and developed as a solution to the limitations of standard RNNs, with a particular focus on issues related to exploding/vanishing gradients in capturing long time dependencies. LSTMs, exploiting a more sophisticated architecture that includes memory cells and gating mechanisms, allow the model to select valuable information and discard overused data. In this way LSTMs are capable of learning both short-term and long-term dependencies between the variables.

The key innovation of LSTM networks lies in the gating mechanism that controls the flow of information through the whole network. An LSTM unit consists of three gates, the forget gate, the input gate, and the output gate, which are used to regulate what information needs to be remembered, what new information needs to be added, and what information should be outputted or deleted. The mathematical operations governing these gates can be summarized as follows:

- **Forget gate**: decides which parts of the previous memory to delete:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

- **Input gate**: Determines which new information to store in the memory:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c)$$

- **Memory cell update**: the new cell state c_t is updated by combining the forget and input gates:
$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$
- **Output gate**: Decides what to output from the current memory state:
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \odot \tanh(c_t)$$

By using these gates, LSTMs are able to maintain the gradient on a stable value, even over very long sequences of information, thus partially solving the vanishing gradient problem that affects standard RNNs. This ability to capture long-term dependencies makes LSTMs particularly well-suited for the task of predicting real estate prices, where historical information, such as past trends or macroeconomic indicators, has a noticeable impact on future prices.

Assumptions and Application to House Price Prediction

In utilizing neural network models, especially RNNs and LSTMs, for house price prediction, we first need to assume that a temporal relationship exists between the variables in our dataset and the output of the model. In our context, house prices are usually temporally dependent: future prices can be contingent upon previous prices and other factors that depend on time. The architecture of RNNs and LSTMs handles exactly this kind of modeling. Furthermore, the model assumes the existence of some non-linear relationships between the data and especially between predictors and the output variable. Neural networks are capable of approximating non-linear functions, which in the context of housing prices can be exploited to capture the complex interactions between the data. For instance, changes in mortgage interest rates, local unemployment rates, or inflation can have non-linear effects on housing prices, which become therefore difficult to model using standard linear regression techniques. In the end, it is postulated that the richness of the dataset used for training the model will be sufficient and representative of general market dynamics. A diverse dataset covering not only historical price information but also a set of economic indicators and local features may significantly improve the prediction of house prices, in particular, within the context of Italian municipalities. These variables are basic to prediction and accurate forecasting as they provide the model with the necessary context to understand how different elements affect house prices over time.

2.2 A Straightforward RNN model

Model structure and architecture

The first dynamic model we test utilizes Long Short-Term Memory (LSTM) layers to predict house prices combining dynamic and static features. The key idea behind this model is to capture both short-term fluctuations and long-term trends in property prices by using LSTM's ability to retain temporal information through its internal memory cells.

Our dataset contains a mixture of dynamic variables that vary over the semester of analysis either at the individual municipal level (e.g., employment rate, average income, etc.) or at the national level (e.g., inflation rate, net domestic product, etc.), and static variables, non-time dependent, at either the municipal level, like the geographical position of a city, its characterization as a mountain or touristic municipality, or at the property level, as the neighbourhood in which a house is located, or its maintenance state.

In this first model, to preserve simplicity, we decide to treat all the features as dynamic, passing them through LSTM layers. In this fashion, static features can be interpreted as non-varying dynamic features, assuming the same value at all time steps. The goal is to identify and learn from latent temporal patterns that might arise from interactions between static features and time-varying factors.

Our RNN-LSTM model is structured as follows:

- **Input Layer**: The input is represented by a sequence of feature vectors containing data of macroeconomic indicators and property/municipality-level characteristics at different time steps (semesters between Jan. 2016 and Dec. 2021). As said, these features include both time-varying factors, such as market prices, economic indicators, and property condition, as well as traditionally static factors, such as property maintenance state, type, and geographic location (neighbourhood, region, area). Even though these static features do not change over time, we treat them as part of the time series, allowing the model to consider them repeatedly at each time step. This, while carrying some drawbacks, should potentially allow the model to explore interactions between features as the housing market evolves.

- **LSTM Layers**: The input sequences are passed through one or more LSTM layers, each of which contains memory cells designed to maintain a hidden state across time steps. LSTM units are equipped with three essential gates:

  - **Forget Gate**: determines which information from the previous time step should be discarded from the cell state. It helps the model forget irrelevant past information that may no longer be useful for future predictions.

  - **Input Gate**: decides which new information from the current time step should be added to the cell state. This allows the LSTM unit to update its internal memory based on new data.

  - **Output Gate**: controls what part of the information should be passed to the next layer or to the output. This allows the LSTM to produce an output that combines both short-term and long-term information.

  These gates enable the LSTM network to retain relevant information across long sequences, addressing the problems related to vanishing/exploding gradient that affect traditional RNNs.

- **Dense Layer**: After passing through the LSTM layers, the sequential data is flattened into a fixed-length vector and passed through a fully connected (dense) layer. This layer processes the information received from the LSTM units to generate a

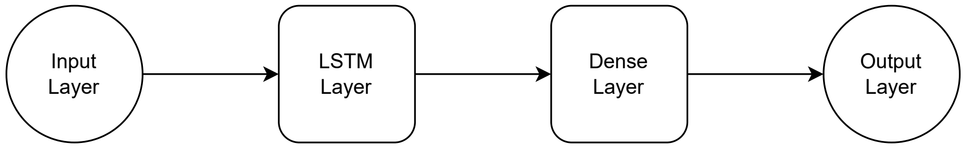

Fig. 1 Recurrent neural network with LSTM layers.

final representation of the input data, combining both dynamic and static features.

- **Output Layer**: The final layer generates the predicted maximum and minimum house prices for the given property at a specific time step. The output is a pair of scalars representing the maximum and minimum house price for the last semester (12th time step), which is derived from the representation learned by the previous layers.

Figure 1 shows a stylized representation of the functioning of a generic recurrent neural network with LSTM layers.

Limitations of the Model

As we discussed above, one of the primary assumptions in this model is that all input features, including static ones, are (and can be) treated as time-varying. This decision is based on the hypothesis that even static features, such as the municipality's location, may exhibit latent interactions with time-varying features. For instance, the desirability of a specific location may evolve over time due to external factors, such as infrastructure development, population changes, or social events.

This approach, of course, while easy to understand and interpret, introduces several limitations. Treating static features as dynamic may result in redundancy or even over-fitting issues. For example, consider a property's geographic location, which is clearly constant over time: the model receives this information repeatedly at every time step incurring in the risk of causing the model to overestimate the importance of this static feature. Additionally, treating static and dynamic features uniformly could obscure important distinctions between time-varying factors (like market prices) and fixed characteristics (like property type or size), potentially leading to inefficiencies in how the model learns from the data. In addition to this, while LSTM networks excel at modeling sequential dependencies, their primary strength lies in predicting outcomes driven by dynamic factors. When static variables dominate the input set, LSTM layers tend to become less effective, resulting in scarce improvements over simpler models, and leading to an unnecessary over-complication of the model and hindering interpretability of results.

2.3 A Double-headed RNN model

Architecture and structure of the model

Considering the limitations of the first model we utilize, with all the inputs being treated as dynamic variables, we now propose a more complex architecture that should possibly address the drawbacks of the previous model, separating static and time-varying inputs. In this way, we allow the model to learn separately from different sets of features, utilizing adequate sequences of layers depending on the nature of the information. The main innovation we introduce is to use a two-headed (two separate input layers) neural network, one to process dynamic features through LSTM layers and the other one to process static features through dense fully-connected layers. In this way we should be able to reduce the risks related to redundancy of information and over-fitting issues, at the same time preserving the ability of the model to learn from temporal patterns exhibited by dynamic features.

The way in which this network is structured is specifically designed to take advantage of the informative power of both sets of features and consists of two separate input pathways that process inputs separately before concatenating them:

- **Double Input Layer**: The input layer of this model consists of two separate branches:
  - **Static Inputs Group**: Features that are static by their nature, such as the geographical location of a municipality, the position of a building, the property type or the maintenance state, are fed to the network into this branch. These features are, indeed, not expected to change over time and thus are processed by a dedicated series of layers. After being inputted, those features are processed in a series of dense layers allowing the network to obtain a fixed high-level representation of the static features. This layers will try to capture the relationships between these variables that do not show any kind of dynamic, time-related interaction between them, solving the issues observed in the previous model.

  - **Dynamic Inputs Group**: The second branch of the input layer contains and handles dynamic features, that vary over time, as property prices, macroeconomic factors, average income of a municipality or its employment rate. These features are then fed into a standard LSTM layer that is responsible for capturing temporal dependencies between the dynamic features, learning patterns from the data sequences, and storing relevant historical information across semesters (time steps).

  The key advantage of this dual-input network lies exactly in its ability to distinguish between static and dynamic features, allowing the model to handle each feature group separately, modeling temporal inter-dependencies between dynamic features while, at the same time, building a static representation of unchanging ones.

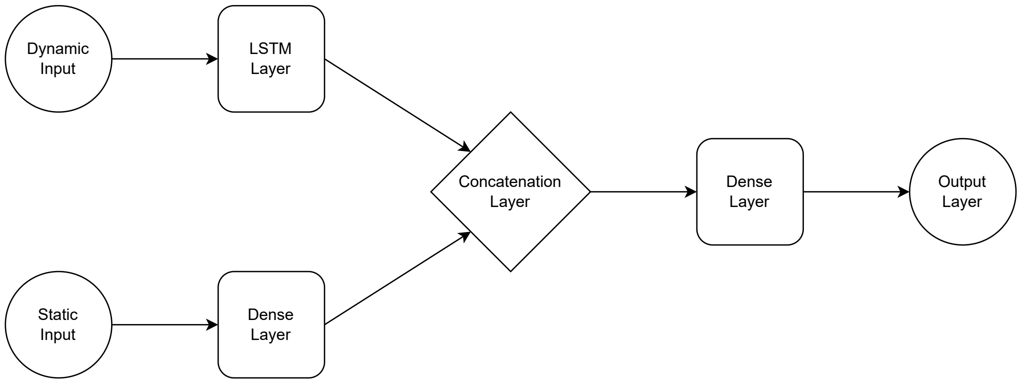

Fig. 2 Two-headed recurrent neural network.

- **Combination Layer**: After the first processing of inputs occurred, all the outputs of the two branches are combined together through a single combination layer, that is used to integrate the temporal patterns learned by the LSTM branch with the high-level representations of static features obtained in the static branch. In this way we are able to obtain a joint output that considers both types of features adequately, providing a more comprehensive understanding of the factors that influence real estate prices.

- **Dense Layers and Output Layer**: The output of the combination layer is then passed through a series of dense layers which are used to aggregate the learned features and refine them for the final prediction, finally capturing eventual interactions between static and dynamic features. As before, the final output of the model is a scalar, representing the predicted maximum and minimum house price for a given property at the last time step of observation (12th semester).

Figure 2 shows a diagram of the structure of our two-headed recurrent neural network.

Strenghts and limitations of the Model

As discussed, one of the main strengths of this model lies in its ability to prevent feature redundancy and over-fitting, avoiding unnecessary repetition of static features over time, creating a more generalized representation of data. Additionally, while the architecture of the network appears clearly as more complex, at the same time it provides a more intuitive and interpretable structure. Indeed, splitting the input into two separate branches, it becomes easier to understand the way in which the two types of features are processed, allowing also for a much finer and flexible tuning of the two branches, modifying separately both the number of nodes and the number of layers before concatenation.

Nonetheless, the model shows also some drawbacks, mainly related to the added complexity of managing two separate input pathways, which increases the number of parameters that need to be learned, making the model more inclined to incur in over-fitting issues when dealing with small datasets, increasing the necessary amount of data. Moreover, the assumption that static features do not change over time, which leads to a separation of inputs, might overlook subtle interactions between static and dynamic features that could exist in certain markets. Another limitation is linked to the risk of losing part of the temporal context when static features are processed exclusively through dense layers: while the combination layer mitigates this to some extent, there may still be scenarios where treating certain static features as, at least partially, dynamic would result in a better performance, capturing the interactions of these features with the dynamic factors.

3 Materials

3.1 Variables of interest and data

The data we use in this study is fully gathered from official governmental sources. The main variable of interest, maximum and minimum prices of real estate, is sourced from the Osservatorio del Mercato Immobiliare (OMI), managed by the Italian revenue authority, Agenzia delle Entrate (AdE), that provides a comprehensive database containing semi-annual updated information regarding both residential and commercial properties. This database is open for free consultation by all the Italian citizens and foreign institutions and individuals, on request, and is widely utilized both for practical and analytical purposes by public and private institutions and individuals. The information provided is based on real estate transaction data, covering all 7 896 Italian municipalities, subdivided in over 27 000 OMI zones, allowing for a precise and localized valuation of properties. These zones can be considered as geographically, economically and socially homogeneous areas, with similar infrastructural conditions, proximity to urban centers, and social development. Considering semi-annual data between the first semester of 2016 and the second semester of 2021, for each municipality, we gather data regarding the recommended price per square meter range for a specific type of property, in a certain zone of the municipality in a given maintenance state. Also, geographical data of the position of the municipality with respect to the Italian peninsula (categorized in: North-West, North-East, Central Italy, Southern Italy, Islands) are provided. Through other official sources, namely, Agenzia delle Entrate, Ministry of Economy and Finance, and the Italian National Institute of Statistics (ISTAT), we also add some relevant information available on an annual basis for each municipality, regarding average income, percentage of inhabitants with an income below 10 000 Euros per year and above 75 000 Euros per year (used as proxies for the lowest and highest wealth levels of a municipality), population density (all the following available on an annual basis for each municipality), employment rate (available on an annual basis for each province), and whether a municipality is considered a touristic area or a mountain area. Last, we added four macroeconomic indicators that can be useful in analysing real estate market fluctuations: inflation rate, Euribor, value of FTSI MIB, Net Domestic Product.

Methodology for Determining Prices

Our variable of interest, the maximum and minimum house prices in a given semester, for a given type of property in a specific area of an Italian municipality, need to be carefully described and understood. The prices provided in the OMI database are not to be considered as direct transaction prices but rather as estimates of an adequate price range for all the properties with certain characteristics located in the area of interest. Those estimates are based on transaction data, but the determination process consists of several steps that involve a robust and stable methodology that incorporates both human judgment and statistical analysis:

1. Each municipality is divided into specific OMI Zones that define contiguous areas characterized by homogeneous characteristics on the social and economic level, allowing for a comprehensive and rational valuation of properties in the area.
2. Market data is collected with a semi-annual frequency and transactions are sampled to obtain, at least, five transactions for each OMI Zone, to provide an adequate amount of information to determine a price range.
3. Using Student-t distribution, a 95% confidence interval around the mean of transaction prices is elaborated, ensuring a robust statistical representation of reality even for smaller municipalities.
4. In some municipalities, where transaction frequency is limited, offer prices may be used instead of final transaction values. In such case, the information is still considered valid but an adjustment coefficient of, at most, 20% is applied to create a more precise alignment with actual market conditions.
5. For areas with insufficient transactions and offer data, virtual valuation methods are applied, by extending data from previous semesters or utilizing, after human expert consultation, prices for similar zones that show both a historical and a socio-economic alignment. In any other case the data is considered insufficient and the municipality receives no valuation.
6. All the observations where the relative deviation between the minimum and maximum price is higher than 0.5, a deeper analysis is performed to verify whether this deviation is caused either by the presence of outliers (that needs to be opportunely excluded), by a great general variability in the data, or by a non-homogeneous sample of properties, especially with respect to the maintenance state of the building (and, consequently, gather more data to elaborate the interval).

Price update over semesters

The OMI evaluation of maximum and minimum prices is performed and updated on a semi-annual basis, taking into consideration both the relevance of newly gathered data and, potential, infrastructural or social changes occurred in the municipalities or in the sub-zones. The update process consists of two distinct procedures, depending on the number of transactions occurred in a specific area:

- **Direct Data Collection**: For municipalities and zones characterized by high transaction frequency, the price is updated on the basis of newly gathered data. The previous evaluation is adjusted by, independently, computing a new price range and then comparing the two intervals, keeping into consideration structural changes, to reflect the most recent market conditions.

- **Indirect Data Collection and Expert Validation**: In areas with lower transaction volumes, the update process relies more heavily on expert assessments and statistical inferences. Local market experts evaluate market changes based on other characteristics and provide a temporary price range that is subsequently evaluated from the technical and statistical point of view by a specialized technical committee, the Comitato Consultivo Tecnico (CCT), to ensure coherence and consistency with broader market trends.

Additionally, the OMI database is subject to partial revisions (other than the fixed semi-annual update) in case of deep changes in market or social conditions, such as the construction of new transportation hubs, strong increase in popularity of a specific location or major infrastructural interventions. The goal is to ensure a good responsiveness of the database to current local and national economic developments.

Potential Drawbacks of OMI Database

The OMI database contains detailed and reliable data regarding price evaluations for all Italian municipalities over the various semesters of the last years. Nonetheless, the usage of these official data involves some drawbacks related mainly to the presence of missing values, especially for areas with limited real estate transaction frequency. The lack of information in those areas can result in multi-period gaps in the data, hindering the possibility of exploiting time-series data efficiently without any form of imputation, or in less precise price estimates. Moreover, the reliance on expert judgments introduces a degree of uncertainty and subjectivity in the evaluations that may alter intrinsic dynamic relationships between variables. Also, the semi-annual updates may generate some lag in the data, without a prompt reaction to actual market changes. Despite the existence of partial, non scheduled, revisions, the database still suffers from the risk of lacking responsiveness to short-term market fluctuations caused, for example, by sudden economic shocks, environmental events or policy-changes.

3.2 Organization of data and descriptive statistics

Dataset Segmentation and characteristics of municipalities

As discussed in depth, municipalities in Italy exhibit substantial variability both in terms of their socio-economic characteristics and property valuation trends: understanding and efficiently exploiting these differences is crucial for a precise valuation of real estate properties through learning-based models.

In order to deal with the deep diversity of the Italian real estate market, the original dataset was divided into three distinct groups, based on the population size of the municipalities:

- **Large Municipalities**: municipalities with more than 50 000 inhabitants; often characterized by a great degree of internal diversity, typically with luxury high-end areas close to the city center, developed infrastructures, and significant flows of tourism. These municipalities see a greater economic activity and should experience, on average, more pronounced short term fluctuations due to economic cycles and authority intervention.

- **Medium-sized Municipalities**: municipalities with a population between 15 000 and 50 000 inhabitants; typically characterized by a mixture of urban and rural areas and elements, with moderate levels of real estate market activities.

- **Small Municipalities**: municipalities with less than 15 000 inhabitants; often located in rural areas, and characterized by a higher level of agricultural activity and land availability. In these municipalities we observe a less active real estate market and low transaction frequency, resulting in more gradual changes in market prices over time.

This segmentation is performed both on practical and analytical grounds, in order to capture more easily the nature of the interaction between our variables while, anyway, preserve a certain degree of heterogeneity between the municipalities. The criteria that we used to divide the municipalities allowed us to capture approximately a third of the Italian population in each segment, creating three uniform groups that can ensure a balanced analysis across varying size of municipalities.

Data Transformation and Pre-processing

To prepare the data for analysis and ensure compatibility with machine learning algorithms, we applied several pre-processing steps. One key transformation is the conversion of categorical variables into dummy variables using one-hot encoding. This method allows categorical data, such as the type of property, property condition, and municipality type, to be represented as binary variables. For example, in the case of the municipality type, each municipality is categorized into a binary column for urban, suburban, and rural categories.

One-hot encoding ensures that the machine learning models can handle categorical features effectively, while also avoiding the pitfalls of treating categorical variables as ordinal (i.e., assigning arbitrary numerical values to categories). By doing this, we improve the model's ability to distinguish between different types of properties and municipalities, capturing the distinct relationships that exist across the groups.

Descriptive Statistics

Descriptive statistics for time-varying features in the three segments are shown in Tables 1-2-3. The income distribution displays an expected increasing trend with the municipality size: the average income in large municipalities is significantly higher than in medium and small ones, reaching approximately €22 616, compared to €21 045 in medium-sized municipalities and €18 647 in small ones. This finding aligns with the general economic theory that larger urban areas tend to concentrate higher-income jobs and economic opportunities.

Table 1 Descriptive statistics for dynamic variables - Big municipalities

Variable	mean	std	min	25% perc.	75% perc.	max
Avg_Inc	22616.066	3838.232	11862.811	20349.687	25189.310	37204.623
Perc_Un10k	28.242	6.529	18.296	23.267	32.041	56.490
Perc_Ov75k	2.873	1.364	0.573	1.900	3.668	8.569
Pop_Density	1293.218	1364.934	51.401	407.179	1776.792	7740.286
Emp_Rate	51.963	17.961	7.937	41.481	66.246	74.050
Min_Price	1325.241	604.811	215.000	900.000	1600.000	6220.000
Max_Price	1786.913	804.373	310.000	1250.000	2150.000	8240.000

Moreover, the percentage of inhabitants earning less than €10 000 per year is higher in small municipalities (31.3 %) than in large ones (28.2 %), reflecting the economic disadvantages of peripheral areas. The inverse pattern is observed for high-income individuals (above €75 000), whose share is minimal in small municipalities (1.2 %) and highest in large ones (2.8 %). Large municipalities exhibit an average density of 1 293 inhabitants per square kilometer, nearly six times that of small municipalities (215 inhabitants per square kilometer). Employment rates appears to be more stable across municipality sizes, with median values consistently around 53 %.

Table 2 Descriptive statistics for dynamic variables - Medium-sized municipalities

Variable	mean	std	min	25% perc.	75% perc.	max
Avg_Inc	21045.973	3816.651	10230.663	18199.864	23815.596	36664.874
Perc_Un10k	28.180	8.275	15.441	21.478	33.791	62.394
Perc_Ov75k	2.056	1.064	0.123	1.232	2.666	9.778
Pop_Density	809.102	865.327	42.533	281.339	989.916	6460.714
Emp_Rate	53.224	18.210	6.817	42.890	67.141	74.050
Min_Price	1143.697	465.292	190.000	800.000	1400.000	3800.000
Max_Price	1523.870	593.257	180.000	1100.000	1850.000	5500.000

Looking at maximum and minimum prices, which represent the dependent variable in our study, the data highlight a positive relationship with municipality size. The average minimum price per square meter is €1 325 in large municipalities, €1 143 in medium-sized ones, and drops to €857 in small municipalities. A similar pattern is observed for maximum prices, with large municipalities reaching an average of €1 786, compared to €1 523 in medium-sized municipalities and €1 134 in small ones. The price range between minimum and maximum values is also larger in big municipalities, indicating a higher variability in property valuations, potentially linked to greater market segmentation.

Table 3 Descriptive statistics for dynamic variables - Small municipalities

Variable	mean	std	min	25% perc.	75% perc.	max
Avg_Inc	18647.577	3889.456	7227.971	15548.512	21505.061	35899.837
Perc_Un10k	31.299	9.724	13.166	23.228	38.652	68.943
Perc_Ov75k	1.247	1.031	0.000	0.488	2.813	7.548
Pop_Density	215.136	292.079	0.747	47.418	250.974	2813.578
Emp_Rate	53.314	17.542	6.817	43.197	66.341	74.050
Min_Price	857.401	460.426	110.000	550.000	1050.000	4900.000
Max_Price	1134.793	601.137	185.000	750.000	1350.000	6300.000

Looking at Table 4, which reports statistics for all municipalities combined, we observe that the aggregation of municipalities masks some of the heterogeneity observed across different municipality sizes. The mean income, for instance, is approximately €19 097, which lies between the figures for small and medium-sized municipalities but significantly underestimates the value for large municipalities. The same applies to the employment rate and population density.

Table 4 Descriptive statistics for dynamic variables - All municipalities

Variable	mean	std	min	25%	75%	max
Avg_Inc	19097.724	4018.861	7227.971	15957.977	22026.312	37204.623
Perc_Un10k	30.791	9.521	13.166	22.934	37.809	68.943
Perc_Ov75k	1.410	1.120	0.000	0.586	2.017	9.978
Pop_Density	330.162	556.467	0.747	57.538	361.762	7740.286
Emp_Rate	53.253	17.645	6.817	43.158	66.486	74.050
Min_Price	910.917	483.634	110.000	570.000	1100.000	6220.000
Max_Price	2108.128	632.731	185.000	780.000	1450.000	8240.000

Table 5 presents the descriptive statistics for the static features. The proportion of municipalities classified as mountain areas is substantially higher in small municipalities (43.5 %), whereas it declines significantly in medium-sized (9.3 %) and large (2.3 %) municipalities. This factor needs to be kept into consideration when performing real estate valuation, as mountainous terrain limits development opportunities and typically results in lower property prices due to accessibility constraints and lower economic activity. Similarly, touristic municipalities are overrepresented among large municipalities (47.7 %), whereas their presence is lower in small municipalities (15.7 %). This aspect suggests a potential driver of price differentials, as touristic appeal tends to sustain higher property values, especially in premium and highly-desired destinations.

A significant distinction is also observed in the type of residential properties. Small municipalities show a higher proportion of villas and cottages (30.7 %), whereas large municipalities feature a greater share of civil houses and economic dwellings. This variation reflects the land availability and housing demand differences between urban and rural areas. The state of conservation of properties appears to be relatively uniform across all municipality sizes, with approximately 15-20% of properties classified as excellent and a negligible

Table 5 Descriptive statistics for static features

Variable	Small (%)	Medium-sized (%)	Large (%)	All (%)
Mountain_City	43.548	9.325	2.342	37.698
Touristic_City	15.772	21.610	47.688	17.696
C	1.486	13.154	21.141	3.688
D	13.204	28.098	26.366	15.572
E	11.336	16.221	19.219	12.246
R	25.540	13.260	9.550	23.396
e_dwellings	26.576	28.257	30.631	26.939
vil_and_cot	30.747	29.073	26.970	30.395
excellent	15.742	20.936	15.976	16.406
poor	0.000	0.000	0.300	0.011
south	25.003	19.234	25.105	24.279
centre	11.404	21.503	22.162	13.078
islands	6.923	9.236	11.231	7.375
north_e	20.339	22.549	20.841	20.637

proportion categorized as poor. This homogeneity suggests that property conditions alone are unlikely to be a major determinant of price differences and that broader economic and locational factors play a more substantial role.

3.3 Data cleaning

As a last step of the data preparation procedures, we need to perform some sort of data cleaning procedures that can allow us to obtain a comprehensive dataset containing data for all the 12 semesters in the analysis. The cleaning process consists of three critical steps: outlier detection and removal, handling administrative changes in municipalities, and managing incomplete sequences in the data, before merging the information of each individual semester into a unique comprehensive dataset.

Outlier Detection and Removal

First, we deal with the presence of outliers in our data. If not properly addressed, outliers can potentially hinder the performance and validity of a predictive algorithm [25, 26] or a statistical test [27, 28] compromising the fidelity of the observed patterns, consequently leading to a potential lack of generalizability and integrity of the resulting models [29]. Given the importance of this task and the wide literature dealing with anomalies detection and removal, we decided to compare the performances of several different outlier detection algorithms to choose the most appropriate for our purposes. We consider eight different outlier detection techniques:

- **Three-sigma rule**: the idea behind this algorithm is that, if an instance falls at a distance farther than three standard deviations from the mean value of a distribution, it has a high probability of being an outlier. Depending on the domain of application, it may also be convenient to use just a 2 or 2.5 sigma distance from the mean to distinguish between in-liers and outliers [30].

- **Inter-quartile Range**: the Inter-quartile range (IQR) is defined as the difference between the third quartile and the first quartile of the distribution. This methodology identifies as outliers the observations lying farther than 1.5 times the IQR from the first quartile (downwards) or the third quartile (upwards).

- **Mean Absolute Distance**: a variation of the three-sigma rule, proposed to overcome some limitations of the two previous techniques linked to distortions in the distribution caused by a heavy presence of outliers. This technique replaces the mean and standard deviation respectively with the median and absolute deviation from the median (MAD), two measures that are not sensitive to anomalies. MAD can be defined as [31]:
$$MAD = E(|x_i - E(x)|)$$

- **Mahalanobis distance**: This measure can be interpreted as a multivariate extension of the classical Z-score, representing a dimensionless covariance distance between data [32]. To compute this measure, let us consider a set of n-dimensional data $x = (x_1, x_2, ..., x_n)^T$, the mean vector $\mu = (\mu_1, \mu_2, ..., \mu_n)^T$ and its covariance matrix Σ. The Mahalanobis distance is expressed as:
$$D(x) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)}$$
All the observations having a distance from the center of the multivariate distribution greater than a set threshold are flagged as outliers.

- **Isolation Forest**: unsupervised methodology that relies on decision trees to identify at an early stage the anomalous observations [33]. Basically, the idea is that the algorithm randomly selects a feature from the dataset and splits the values at a random point between the maximum and the minimum value and iterates this procedure until each branch contains a single data point or a certain path length is reached. Subsequently, after creating a series of random isolation trees, the expected path length needed to isolate a certain point is computed and an anomaly score is calculated for each point. Then, an anomaly threshold is defined and all the data points showing a score that violates such threshold are marked as outliers [34, 35].

- **Density-Based Spatial Clustering of Applications with Noise**: The algorithm considers a certain target distance ϵ and a minimum number of points m: considering a single data point, we define "neighbour" any point with a distance smaller than ϵ from the original chosen point. When the number of neighbours of a point is greater than m, the group of points is considered a cluster. The algorithm then defines as core points those that have at least m neighbours, as border points those that cannot be considered core points but have a core neighbour, and as outliers those that do not belong to any other category.

- **One-Class Support Vector Machines**: another unsupervised anomaly detection technique. The idea behind this technique is to model the underlying distribution of data while being insensitive to anomalies in the training records. The input space is implicitly mapped by a kernel function to a higher-order dimensional space so that a clearer distinction between normal and outlying data is possible [36]. The algorithm indeed will take data and project them into a higher dimensional space, using a certain transformation function, and then identify an optimal decision boundary, separating normal points from outliers.

- **Local Outlier Factor**: a density based technique, particularly well suited to identify local outliers as it considers the density of a point comparing it with the density of its neighbours. A LOF score is computed, for each point of a distribution, as the ratio between the average density of the neighbourhood to the density of the point. If the score is greater than one (even though a tolerance margin can be set) the point is more likely to be marked as an outlier [37].

To choose the most appropriate outlier detection technique, among the ones mentioned above, we compare the results obtained with the different algorithms, based on all the three datasets we considered. Each algorithm is used to clean the raw data and the anomalies found are compared. Specifically, we are interested in monitoring the percentage of outliers that each algorithm identifies for each period of analysis in each dataset. Our goal is to find a technique that is able to identify the outliers without deleting an excessive amount of observations, compared to the other algorithms. In a successive set of analysis we also observe whether the various algorithms identify the same observations as anomalies. Last, we monitor how spread the anomalies are with respect to the different municipalities in Italy. Indeed, univariate methodologies, that focus only on the distribution of prices, suffer from the risk of marking as outliers all these observations coming from properties located, for example, in high-level luxury municipalities, without considering other characteristics as the average income or level of wealth of inhabitants. In our context, anomalies should rather be interpreted as mis-priced properties that do not have any characteristic able to justify the difference in price with respect to similar properties in comparable locations.

After carefully evaluating the results of the comparison, we decide to proceed with the algorithm that exploits One-Class Support Vector Machines to identify and remove anomalous observations, as the best suited for the data at hand.

Handling Municipalities Changes

Another issue that needs to be considered is related to the noticeable number of administrative changes that Italian municipalities have undergone in the last years, and in particular in the time period considered. Several municipalities have indeed been subject to name changes, municipal mergers or splits, and changes of province. According to ISTAT, between 2011 and 2024, the number of municipalities decreased from 8 092 to 7 896, with mergers and splits. These modifications primarily involved small, neighboring municipalities consolidating for administrative efficiency, but they represent a non-trivial issue for our analysis, as data of these municipalities need to be gathered and adequately managed. Utilizing official data from ISTAT and from the Dipartimento per gli Affari Interni e Territoriali, we systematically reconstructed the data sequence for each municipality, utilizing the updated list of names, provinces and existing municipalities as of the last semester of the considered time period (second semester of 2021).

Handling Incomplete Sequences

We utilize data for the semesters between 2016 and 2021 (12 semesters in total). As discussed before, time sequences may have some data gaps, especially for small municipalities with limited real estate transaction frequency. RNNs need complete data to properly learn the complex relationships existing between the variables. After careful consideration, we decided, rather than using imputation methods to fill gaps of unavailable data, which could introduce biases and uncertainty, to keep only data of municipalities with complete data sequences across all 12 periods of observation, to ensure robustness of the models and reliability of the whole set of data fed into them.

4 An empirical investigation on Italian municipalities

We now present the results of our study on the application of RNNs to real estate price prediction and valuation in Italy. Specifically, we intend to evaluate the ability of Recurrent Neural Networks to forecast real estate prices. Our idea is to exploit and capture temporal relationships between our predictors to predict the maximum and minimum price per square meter for a certain type of residential building in a specific area, integrating in the analysis also macroeconomic and financial factors, at the national, regional and municipal level. These factors include macroeconomic indicators (inflation rate, Euribor, net domestic product, and the FTSE MIB index), local socioeconomic indicators (average income levels, employment rates, and population density), and past observed prices. We consider an observation window of 12 semesters, between the first semester of 2016 and the last semester of 2021, including data from 7 764 Italian municipalities. Our models exploit the prices and other data of the first 11 semesters and tries to forecast the prices in the 12th semester. As said, municipalities were segmented into three groups to account for the high degree of structural heterogeneity of the Italian real estate market, allowing for a clearer assessment of the model performance across different economic and demographic environments.

The results presented in the following subsections include a comparative evaluation of the predictive accuracy of two distinct RNN architectures: a standard RNN, which treats all input features as dynamic, and a double-input layer RNN, which processes dynamic and static features separately before combining them in a later stage of the network. The model performance is assessed using standard error metrics (Root Mean Square Error, Mean Absolute Error, and Mean Absolute Percentage Error) both in-sample and out-of-sample. The optimal configuration of each model, including the number of neurons per layer, activation functions, and regularization strategies, has been determined through an extensive hyperparameter tuning process.

The Standard RNN consists of a single LSTM layer, followed by a fully connected dense layer and an output layer. After training the network and evaluating the performances with various combinations of layers and activation functions, we decide to set the number of LSTM units in the hidden layer to 29 neurons.

The Double-Input Layer RNN is divided into two groups: the dynamic input branch consists of a single LSTM layer with 30 units, similar to the Standard RNN. Conversely, the static input branch is composed of two dense layers with 32 and 34 neurons, respectively, to allow the model to extract representations from the static attributes. The outputs of the two groups are concatenated and passed through an additional dense layer with 28 neurons before reaching the output layer. Additionally, we employ the $ReLU$ activation function whereas the LSTM units utilize the standard *tanh* activation function for efficient gate computations.

To evaluate the model generalization we implement a *K-fold cross-validation* procedure with $K = 10$. This technique partitions the dataset into ten approximately-equally sized subsets, iteratively training the model on nine subsets while reserving the remaining subset for validation. This process is repeated ten times, ensuring that each observation is used for both training and validation, thereby reducing the sensitivity of performance metrics to the specific train-test split. Moreover, to account for potential randomness in weight initialization and gradient updates, we repeat the training procedure for each fold using 20 different random initializations, averaging the results to obtain robust performance estimates. This methodology allows us to minimize overfitting risks and to assess the consistency of our models across different subsets of the dataset.

Given the complexity of the models, we also utilize the *L2 regularization* with a penalty coefficient $\lambda = 0.001$ to the weight parameters of all dense layers. This penalization term discourages excessively large weight values, preventing overfitting while maintaining model flexibility. Beyond this, we also compute average partial derivatives and the Horel-Giesecke significance statistics of the neural network outputs with respect to the input variables, aiming to assess the relative importance of each feature in the determination of the price range of an observation unit. While static characteristics such as geographic location and structural attributes remain critical for real estate valuation, our analysis focuses on the explanatory power of time-varying features. This choice is also motivated on the open debate that exists on the evaluation of the relevance of inputs represented by dummy variables that are not continuous by definition and thus cannot take any value between 0 and 1, making the interpretation of a perturbation to compute partial derivative less intuitive and clear.

4.1 Forecasting maximum and minimum prices

The empirical analysis presented herein evaluates the predictive performances of the two recurrent neural network architectures across municipalities segmented by population size, as shown in Tables 6, 7, and 8. The outcomes underline that the effectiveness of each model varies noticeably depending upon the segment considered, possibly due to differences in the structural characteristics of the municipalities considered.

For large municipalities (Table 6), typically characterized by complex market dynamics and high transaction frequency, the Double-Input Layer RNN consistently outperforms the standard RNN. Specifically, the double-input architecture shows a reduction in RMSE, MAE, and MAPE values both in-sample and out-of-sample, indicating its capability in capturing latent temporal dependencies between our variables. The improved predictive accuracy could be attributed to the dedicated treatment of static and dynamic variables separately, allowing for more precise modeling of interactions typical in larger markets.

Table 6 Performances of the two models in the large municipalities segment

	Standard RNN		Double-Input Layer RNN	
	in-sample	out-of-sample	in-sample	out-of-sample
RMSE	150.367	164.580	145.677	154.119
MAE	112.524	122.057	105.782	110.392
MAPE	0.086	0.097	0.077	0.085

In medium-sized municipalities (Table 7), the relative superiority of the Double-Input Layer RNN becomes even more evident as the RMSE, MAE, and MAPE show significant improvements compared to the Standard RNN, suggesting that the differentiation in input handling substantially improves the predictive performance also in environments characterized by moderate complexity and heterogeneity. Such municipalities likely exhibit market dynamics driven by a mix of stable location-based characteristics and gradually varying socio-economic factors, both effectively handled by the dual-input structure of the model.

Table 7 Performances of the models in the medium-sized municipalities segment

	Standard RNN		Double-Input Layer RNN	
	in-sample	out-of-sample	in-sample	out-of-sample
RMSE	134.393	145.650	116.266	123.520
MAE	102.831	110.074	85.400	89.522
MAPE	0.081	0.084	0.075	0.076

Conversely, small municipalities (Table 8), which are typically less dynamic due to lower transaction frequency and fewer market fluctuations, demonstrate the opposite scenario. Here, the Standard RNN model achieves slightly better performances in our metrics, compared to the Double-Input architecture. This suggests that the separate handling of static and dynamic inputs in the Double-Input model, while beneficial in complex contexts, might introduce unnecessary complexity in more stable environments, resulting in minor but noticeable performance losses compared to the simpler traditional structure.

Table 8 Performances of the two models in the small municipalities segment

	Standard RNN		Double-Input Layer RNN	
	in-sample	out-of-sample	in-sample	out-of-sample
RMSE	89.203	94.648	98.623	99.372
MAE	59.230	62.329	64.908	66.031
MAPE	0.064	0.065	0.072	0.072

4.2 Analyzing input relevance

Tables 9, 10, and 11 report the analyses of input relevance based on the average partial derivatives (PD) and the Horel-Giesecke significance statistics (HF).

For big municipalities, Table 9 highlights that historical prices (minimum and maximum prices from previous observations) dominate the input relevance measures, highlighting some sort of market inertia and stability in urban environments. Beyond past prices, average income levels and macroeconomic indicators such as the Euribor rate and the national domestic product (NDP) significantly impact price dynamics. This reflects a direct link between the purchasing power of residents, borrowing costs, and overall economic health on property valuation in urban settings, where market conditions are particularly sensitive to broader economic fluctuations.

Table 9 Evaluation of input relevance for large municipalities

Variable	PD_min	HF_min	PD_max	HF_max
Min_Price	2470.8	61969.7	3308.8	111076.9
Max_Price	2456.5	61223.5	3329.7	112539.7
Avg_Inc	923.1	8694.6	1243.0	15763.3
Euribor12	672.2	4646.2	905.4	8428.6
Perc_Un10k	662.2	4500.3	892.9	8180.8
Italy_NDP	649.7	4316.9	874.2	7817.2
Emp_Rate	499.6	2571.5	672.4	4657.8
FTSE_MIB	-471.1	2242.8	-634.6	4069.8
Perc_Ov75k	-425.1	1825.1	-572.5	3310.6
InflationRate	80.0	85.4	108.4	156.4
Pop_Density	55.6	37.3	73.9	66.3

Medium-sized municipalities (Table 10) present a slightly different scenario. While historical prices remain crucial predictors, macroeconomic indicators emerge with increased significance compared to large municipalities. Specifically, variables such as inflation rates and Euribor prominently influence the real estate prices, as evidenced by their PD and HF statistics in the analysis. This heightened sensitivity indicates that also medium-sized municipalities' property markets may react substantially to changes in national economic conditions like larger, and more resilient, markets.

Table 10 Evaluation of input relevance for medium-sized municipalities

Variable	PD_min	HF_min	PD_max	HF_max
Min_Price	1696.8	29180.8	2028.3	41709.2
Max_Price	1500.0	22811.1	2289.7	53134.3
Euribor12	711.1	5125.4	957.5	9291.6
InflationRate	579.7	3431.3	786.5	6313.8
FTSE_MIB	437.0	1942.3	591.0	3552.6
Emp_Rate	339.0	1190.1	455.5	2149.8
Avg_Inc	234.9	568.3	321.2	1061.7
Perc_Un10k	232.7	551.6	317.6	1022.2
Pop_Density	103.9	115.0	143.4	218.5
Perc_Ov75k	-81.6	68.0	-113.9	132.5
Italy_NDP	38.5	17.5	56.2	36.6

Small municipalities (Table 11) exhibit distinct characteristics, with even greater reliance on broader macroeconomic indicators. Besides the considerable influence of past maximum and minimum prices, the variables such as Euribor, inflation rate, and Italy's national domestic product gain relative importance, suggesting that smaller, less active real estate markets are highly responsive to macroeconomic fluctuations. This can be attributed to lower local economic diversification, making these municipalities more vulnerable to external economic shocks and macroeconomic cycles.

5 Conclusions

We explored the possibility of utilization of Recurrent Neural Networks (RNNs) for the prediction and evaluation of real estate prices in Italian municipalities. Testing the robustness of our models across three distinct datasets comprising properties located in municipalities of different sizes, we show the ability of RNNs with Long Short-Term Memory (LSTM) architectures to capture the hidden non-linear and temporal relationships between the prices and our predictors.

With respect to the results of the predictive task, the traditional RNN model slightly outperformed the double-input model in the small municipalities, characterized by an overall reduced financial activity and transaction frequency, while the double-input model proved superior for small and large municipalities, characterized by a higher level of complexity both in the market dynamics and in the price determination. Both models, anyway, obtain satisfactory results, achieving, on average, a prediction error consistently below the 10% threshold.

On the explanatory side, we find that, for all the three datasets, previous observations of maximum and minimum prices remain the strongest predictors of the prices in the last observation period, suggesting a certain degree of stability in the maximum and minimum prices per sqm of real estate properties in the municipalities. The sensitivity of the models to the other dynamic variables, on the other hand, varies depending on the size of the municipality analyzed, showing, anyway, a consistent importance of macro-economic indicators such as inflation rate and net domestic product.

Overall, our analysis shows the potential of RNN-based approaches for improving the prediction of real estate valuations, which becomes particularly relevant in contexts characterized by high variability and strong temporal dependencies.

While we payed particular attention to the methodological side of the study and to the quality of data, using official data for almost the whole Italian territory, our approach is not free from limitations, such as the exclusion of municipalities with incomplete data sequences, which could reduce the generalizability of the results. Moreover, the distinction between static and dynamic variables, while on the one hand may have reduced redundancy and overfitting-related risks, on the other hand it might also have masked subtle interactions between these two types of features. Last, our reliance on semiannual updates of data may have influenced the inherent variability of data that could need more frequent updates for bigger municipalities with higher transaction frequency while resulting even exaggerated for very small municipalities where properties may also remain untraded for years.

Future research should focus on benchmarking these models against other models traditionally used for these tasks or against other advanced architectures, such as attention-based LSTM networks or transformers. Furthermore, experimenting with different hybrid architectures that integrate convolutional layers or graph-based methods could provide additional flexibility in modeling spatial relationships and market dynamics. Finally, the applicability of these methods to other countries or regions with distinct housing market characteristics should be explored, providing a broader testbed for validating and refining the proposed models.

References

[1] Case, K.E., Shiller, R.J.: Is there a bubble in the housing market? Brookings Papers on Economic Activity **2003**(2), 299–362 (2003)

[2] Bourassa, S.C., Hoesli, M., Scognamiglio, D.: Housing finance, prices, and tenure in switzerland. Swiss Journal of Economics and Statistics **147**(3), 225–245 (2011)

[3] Rosen, S.: Hedonic prices and implicit markets: Product differentiation in pure competition. Journal of Political Economy **82**(1), 34–55 (1974)

[4] Crawford, G.W., Fratantoni, M.C.: Assessing the forecasting performance of regime-switching, arima and garch models of house prices. Real Estate Economics **31**(2), 223–243 (2003)

[5] Stevenson, S.: A comparison of the forecasting ability of arima models. Journal of Property Investment & Finance **25**(3), 223–240 (2007)

[6] Hill, R.C., Knight, J.R., Sirmans, C.F.: Estimating capital asset price indexes. Review of Economics and Statistics **81**(2), 226–234 (1999)

[7] Brooks, C.: Introductory Econometrics for Finance. Cambridge University Press, Cambridge, MA (2014)

[8] Makridakis, S., Spiliotis, E., Assimakopoulos, V.: Statistical and machine learning forecasting methods: Concerns and ways forward. PLOS ONE **13**(3), 0194889 (2018)

[9] Nilsson, P.: Prediction of residential real estate selling prices using neural networks. Retrieved from https://www.diva-portal.org/smash/get/diva2:1304961/FULLTEXT01.pdf (2019)

[10] Zhang, G., Patuwo, E.B., Hu, M.Y.: Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting **14**(1), 35–62 (1998)

[11] Ćetković, J., Lakić, S., Lazarevska, M., Žarković, M., Vujošević, S., Cvijović, J., Gogić, M.: Assessment of the real estate market value in the european market by artificial neural networks application. Complexity **2018**(1), 1472957 (2018)

[12] Chiarazzo, V., Caggiani, L., Marinelli, M., Ottomanelli, M.: A neural network based model for real estate price estimation considering environmental quality of property location. Transportation Research Procedia **3**, 810–817 (2014)

[13] Peterson, S., Flanagan, A.: Neural network hedonic pricing models in mass real estate appraisal. Journal of real estate research **31**(2), 147–164 (2009)

[14] Worzala, E., Lenk, M., Silva, A.: An exploration of neural networks and its application to real estate valuation. Journal of Real Estate Research **10**(2), 185–201 (1995)

[15] Duca, J.V., Muellbauer, J., Murphy, A.: What drives house price cycles? international experience and policy issues. Journal of Economic Literature **59**(3), 773–864 (2021)

[16] Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge, MA (2016)

[17] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997)

[18] Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. IEEE

transactions on Signal Processing **45**(11), 2673–2681 (1997)

[19] Nelson, D.M., Pereira, A.C., Oliveira, R.A.: Stock market's price movement prediction with lstm neural networks. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN) (2017). IEEE

[20] Fischer, T., Krauss, C.: Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research **270**(2), 654–669 (2018)

[21] Kim, S., Kang, M.: Financial series prediction using attention lstm. arXiv preprint arXiv:1902.10877 (2019)

[22] Peng, H., Li, J., Wang, Z., Yang, R., Liu, M.: Lifelong property price prediction: A case study for the Toronto real estate market (2021)

[23] Bin, J., Tang, S., Liu, Y., Wang, G., Gardiner, B., Liu, Z., Li, E.: Regression model for appraisal of real estate using recurrent neural network and boosting tree. In: 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), pp. 209–213 (2017). IEEE

[24] Zhao, Y., Chetty, G., Tran, D.: Deep learning with xgboost for real estate appraisal. In: IEEE Symposium Series on Computational Intelligence, pp. 1884–1891 (2019). IEEE

[25] Ardeti, V.A., Kolluru, V.R., Varghese, G.T., Patjoshi, R.K.: An outlier detection and feature ranking based ensemble learning for ecg analysis. International Journal of Advanced Computer Science and Applications **13** (2022) https://doi.org/10.14569/IJACSA.2022.0130686

[26] Choi, S.W.: The effect of outliers on regression analysis:regime type and foreign direct investment. Quarterly Journal of Political Science **4** (2009) https://doi.org/10.1561/100.00008021

[27] Zimmerman, D.W.: A note on the influence of outliers on parametric and non-parametric tests. Journal of General Psychology **121** (1994) https://doi.org/10.1080/00221309.1994.9921213

[28] Zimmerman, D.W.: Increasing the power of nonparametric tests by detecting and downweighting outliers. Journal of Experimental Education **64** (1995) https://doi.org/10.1080/00220973.1995.9943796

[29] Ozsahin, D.U., Mustapha, M.T., Mubarak, A.S., Ameen, Z.S., Uzun, B.: Impact of outliers and dimensionality reduction on the performance of predictive models for medical disease diagnosis. (2022). https://doi.org/10.1109/AIE57029.2022.00023

[30] Miller, J.: Reaction time analysis with outlier exclusion: Bias varies with sample size. The Quarterly Journal of Experimental Psychology Section A **43**(4), 907–912 (1991)

[31] Huber, P.J.: Robust Statistics vol. 523. John Wiley & Sons, Hoboken, NJ (2004)

[32] Li, X., Deng, S., Li, L., Jiang, Y.: Outlier detection based on robust mahalanobis

distance and its application. Open Journal of Statistics **9**(1), 15–26 (2019)

[33] Liu, F.T., Ting, K.M., Zhou, Z.-H.: Isolation forest. In: 2008 Eighth Ieee International Conference on Data Mining, pp. 413–422 (2008). IEEE

[34] Lesouple, J., Baudoin, C., Spigai, M., Tourneret, J.-Y.: Generalized isolation forest for anomaly detection. Pattern Recognition Letters **149**, 109–119 (2021)

[35] Xu, D., Wang, Y., Meng, Y., Zhang, Z.: An improved data anomaly detection method based on isolation forest. In: 2017 10th International Symposium on Computational Intelligence and Design (ISCID), vol. 2, pp. 287–291 (2017). IEEE

[36] Erfani, S.M., Rajasegarar, S., Karunasekera, S., Leckie, C.: High-dimensional and large-scale anomaly detection using a linear one-class svm with deep learning. Pattern Recognition **58**, 121–134 (2016)

[37] Xu, H., Zhang, L., Li, P., Zhu, F.: Outlier detection algorithm based on k-nearest neighbors-local outlier factor. Journal of Algorithms & Computational Technology **16**, 17483026221078111 (2022)