



## IPS Abstract

### **Applied Machine Learning for Central Bank Statistics: Supervised Models for the detection of Social Housing Complexes in Chile**

**Author:** Mr Martin Rebolledo

**Submission ID:** 1150

**Reference Number:** 1150

#### **Brief Description**

Each trimester, the Central Bank of Chile calculates the Household Price Index.

This publication serves as an indicator for the price trends of the country's housing market.

The index is based on administrative records of housing transactions, which include various types of properties.

To ensure the accuracy of the index, transactions that are part of fully subsidized social housing complexes must be identified and excluded.

This paper addresses this need by proposing a supervised classification approach to identify social housing records as a binary statistical classification problem using Machine Learning Models.

The chosen model, an Artificial Neural Network (ANN) with dropout regularization to prevent overfitting, achieves a high rate of accuracy in detecting social housing transactions.

Results show that from 2004 to 2022, approximately 8% of all transactions in Chile corresponded to properties under fully subsidized programs.

Furthermore, our study highlights the potential of machine learning for automating data processing in Central Banks, which can lead to enhanced accuracy in the creation of official statistics.

#### **Abstract**

Each trimester, the Central Bank of Chile calculates the Household Price Index. This publication serves as an indicator for the price trends of the country's housing market. The index is based on administrative records of housing transactions, which include various types of properties. To ensure the accuracy of the index, transactions that are part of fully subsidized social housing complexes must be identified and excluded. This paper addresses this need by proposing a supervised classification approach to identify social housing records as a binary statistical classification problem using Machine Learning Models. The chosen model, an Artificial Neural Network (ANN) with dropout regularization to prevent overfitting, achieves a high rate of accuracy in detecting social housing transactions. Results show that from 2004 to 2022, 8% of all transactions in Chile corresponded to social properties. Furthermore, our study highlights the potential of machine learning for automating data processing in Central Banks, which can lead to enhanced accuracy in the creation of official statistics.