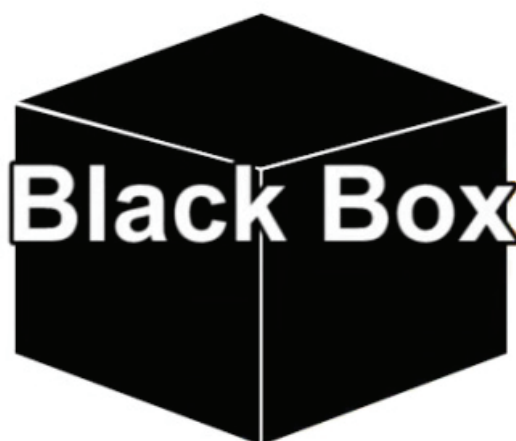


Redes Neurais Artificiais: um Estudo de Caso no Mercado Imobiliário de Castelo Branco

Maria Cristina Canavarro Teixeira

Escola Superior Agrária, Quinta da Sra. de Mércules,
6001-909 Castelo Branco, Portugal
ccanavarro@ipcb.pt



Sumário

As redes neuronais artificiais (RNA's) são uma ferramenta de modelação poderosa com aplicações em diversas áreas de conhecimento como por exemplo na Medicina, nos Mercados Financeiros e na Indústria. Os modelos de RNA's têm vindo a ser aplicados ao longo do tempo, experimentando novos algoritmos e diferentes arquitecturas, sempre com o objectivo de se encontrar um modelo que reproduza a realidade o mais fielmente possível. Uma investigação desenvolvida na cidade de Castelo Branco - Portugal, mostrou que com as RNA's se consegue um desempenho melhor, na ordem dos 10%, do que com as metodologias tradicionais usadas na estimação do preço da habitação (metodologias hedónicas).

Palavras Chave: Redes Neurais Artificiais; Preço da habitação; Castelo Branco – Portugal.

Resumo

O cérebro humano é o sistema de cálculo mais completo que o homem conhece. A capacidade que o cérebro humano tem, de pensar, recordar e resolver problemas, inspirou muitos investigadores, na tentativa de procurar modelar no computador o funcionamento do cérebro humano.

A reflexão sobre a consciência, e qual o órgão que a contém, dura desde acerca de dois mil anos, com os filósofos gregos a serem dos primeiros a especular sobre a localização da alma. O conhecimento sobre como o cérebro funciona é resultado de uma investigação feita nos últimos 100 anos. Ramón y Cajal¹ em 1894 foi o primeiro a propor a teoria dos neurónios (Cajal, 1990). Desde então, numerosos foram os progressos na compreensão do cérebro humano e dos constituintes dos neurónios naturais: os axónios, as dendrites, as sinapses e as activações electro-químicas (Rojas, 1996).

Uma rede neuronal artificial (RNA) é um modelo matemático inspirado em redes neuronais biológicas.

As RNA's são constituídas por um conjunto de neurónios com capacidade de processamento. Estes neurónios estão ligados entre si, para que a informação circule entre todos eles, através de impulsos (simulando o cérebro humano) que são desencadeados por funções matemáticas. A maior parte das RNA's são sistemas adaptativos que alteram a sua estrutura com base em informações internas e externas que circulam pela rede neuronal durante a fase de treino da rede. É por este motivo, que estes modelos são do tipo “caixa-negra” porque não se “vê” o seu funcionamento interno.

A primeira versão de um neurónio artificial foi proposta por McCulloch e por Pitts em 1943 e mereceu o nome Psychon (McCulloch e Pitts, 1943). Nesta versão, existe um conjunto de entradas (que simulam as dendrites) que têm associados pesos (que simulam a intensidade e o carácter de excitação ou de inibição, isto é, o impulso). É calculada a soma pesada das entradas que, se ultrapassar um valor limiar de referência, faz com que o neurónio dispare, produzindo um sinal à sua saída, tal como mostra a Figura 1.

Depois de avanços e recuos, a comunidade científica conseguiu devolver o interesse e a confiança, no fascinante campo da computação neuronal. Actualmente existem vários tipos de redes de neurónios artificiais, tendo sido propostas ao longo do tempo várias classificações. Uma classificação possível envolve três aspectos: o modo como os neurónios

¹ Ramón y Cajal (1852-1934). Médico e histologista espanhol, conhecido pelo seu trabalho sobre o cérebro e nervos, isolando o neurónio e descobrindo como os impulsos nervosos são transmitidos às células do cérebro.

se ligam entre si (topologia da rede), como ficam activos ou inactivos (função matemática de activação) e a forma como modificam alguns dos seus elementos (dinâmica da rede).

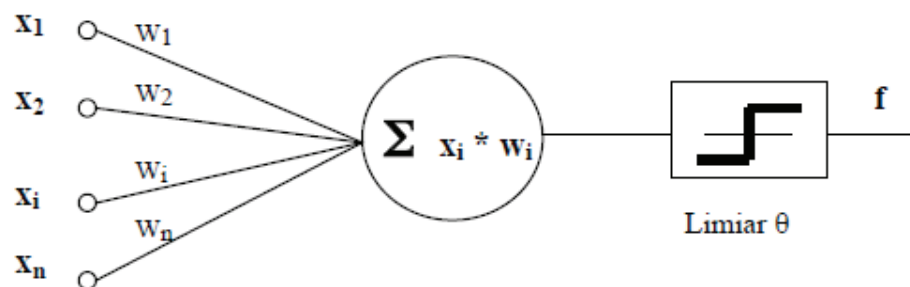


Figura 1: Neurónio artificial segundo McCulloch e Pitts (McCulloch e Pitts, 1943)

A aplicação das redes neuronais artificiais em Portugal, já data de alguns anos a esta parte. Desde as áreas das tecnologias informáticas, às energias renováveis, à sismologia e até mesmo à previsão da qualidade da carne, são variadíssimas as aplicações desta metodologia na investigação realizada em Portugal. Inclusive, em 2009, dois jovens investigadores portugueses da Faculdade de Ciências e Tecnologia da Universidade de Coimbra (FCTUC), venceram o concurso anual internacional, promovido pela PhysioNet, organização do MIT (Instituto Tecnológico de Massachusetts) que gere uma das maiores bases de dados clínicos do mundo para fins de investigação. A aplicação informática vencedora, assenta numa arquitectura avançada de redes neuronais artificiais e foi considerada “o melhor algoritmo computacional para previsão de episódios de hipotensão arterial”².

Os métodos baseados na Inteligência Artificial (IA), aplicados à estimação dos preços dos imóveis, começaram a ter um interesse crescente na década passada, embora a aplicação destes métodos no mercado imobiliário Português se encontre ainda em estado embrionário.

Existe porém, uma metodologia muito usada para estimar o preço dos imóveis em função das suas características – os modelos de preços hedónicos – que são modelos de regressão linear múltipla (Rosen, 1974). Estes modelos têm sido desenvolvidos em vários estudos para o mercado imobiliário Português, assim como para outros locais do mundo e têm demonstrado ser bastante eficazes como pode ler-se em Canavarro et al. (2010a).

Com o objectivo de conhecer e entender o mecanismo de formação do preço da habitação para a cidade de Castelo Branco, recolheu-se uma amostra representativa dos apartamentos vendidos na cidade entre 2005 e 2009,

² Conferência Computers in Cardiology 2009 (<http://cinc2009.org/>)

através da mediação imobiliária de quatro empresas sediadas em Castelo Branco. A LING, a IMOFATOR, a SGH e a GRADUZ através dos seus responsáveis, disponibilizaram-se no fornecimento das características e do preço dos apartamentos por eles vendidos. Foi possível a constituição de uma base de dados com mais de duas centenas de apartamentos e trinta e quatro características, quer de natureza quantitativa quer qualitativa. Após um estudo exaustivo destas características, foi identificado um conjunto de variáveis como sendo as mais determinantes na explicação do preço. Foram usadas técnicas estatísticas na determinação dessas variáveis, que posteriormente também foram validadas pelos Agentes Imobiliários, para que o modelo representasse o mais fielmente possível, a realidade do mercado imobiliário em Castelo Branco.

Seguiu-se a estimação do modelo neuronal para o preço de venda do apartamento, em função das variáveis previamente identificadas, a saber, a área útil (metros quadrados), a localização (geográfica dentro da cidade), o índice de conforto e a existência de anexos (garagem e arrecadação). No caso de apartamentos usados o modelo inclui também o estado de conservação do apartamento e o ano em que ocorreu a venda. O facto de, durante o período de anos amostrado o preço dos imóveis ter sofrido um decréscimo, principalmente os usados, revelou ser um factor importante na determinação do preço. Na seguinte equação escrevemos de forma simplificada este modelo.

$$\text{Preço} = b_0 + b_1 \text{Área útil} + b_2 \text{IConforto} + b_3 \text{IAnexos} + b_4 \text{ILoc} + b_5 t \square \text{Estado} + b_6 \text{IConserv} \square \text{Estado} + e$$

Na estimação do modelo, optou-se pelo uso do software SPSS v.17³, embora exista outro tipo de software específico de RNA's como por exemplo o Neuroshell⁴ ou o NeuroSolutions⁵. O desenho da rede estimada, assim como o gráfico de dispersão dos valores do preço estimado pela RNA vs os valores dos preços observados, podem ser observados na Figura 2.

No desenho da rede podemos ver a sua estrutura, com uma camada de neurónios de entrada (com a informação das variáveis de entrada, ou independentes), uma camada intermédia com quatro neurónios e um único nó final, que devolve o preço do apartamento vendido.

Em relação ao ajuste do modelo, o SPSS devolve uma medida de erro, a partir da qual podemos calcular a eficiência da rede. Nesta rede, obteve-se uma eficiência de aproximadamente 89%, podendo visualizar-se o bom ajuste através do gráfico de dispersão (Figura 2).

³ <http://www.spss.com>

⁴ <http://www.neuroshell.com>

⁵ <http://www.neurosolutions.com>

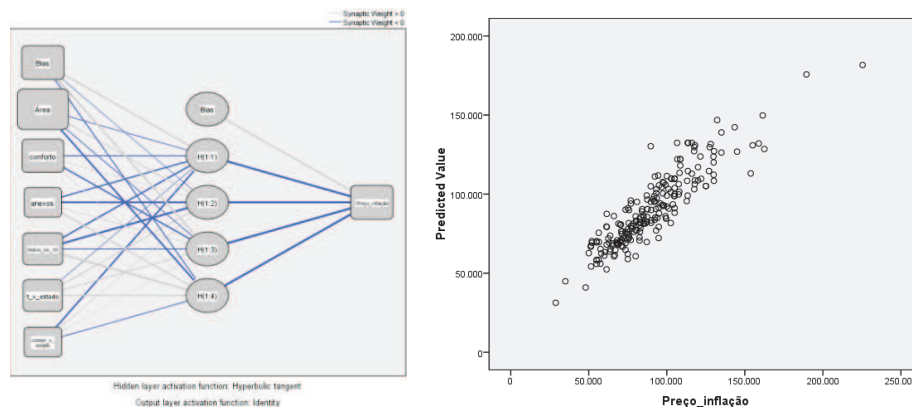


Figura 2: Estimação da RNA para o preço de um apartamento (saída do SPSS v. 17, opção Neural Networks (MLP))

Com o objectivo de comparar a precisão do modelo neuronal com as tradicionais metodologias hedónicas, foram calculadas outras medidas de precisão, tendo sido possível constatar a superioridade do modelo neuronal em relação aos tradicionais modelos hedónicos (Canavarro et al., 2010b). De facto, encontram-se na literatura, resultados que permitem afirmar que no caso de dados de Mercado Imobiliário, os erros dos modelos de IA variam entre 5% e 10%, enquanto que os modelos de regressão linear múltipla, variam entre 10% e 15%.

Outra vantagem de um sistema de IA é a sua capacidade para estimar o valor das propriedades que apresentam características significativamente diferentes das que estão nas proximidades (valores extremos ou outliers), dado que este tipo de sistemas submete as amostras a processos matemáticos muito mais complexos que o modelo de Regressão Múltipla.

Uma questão se coloca: porquê usar as RNA's? E uma resposta pode ser dada; porque têm uma apetência natural em adquirir o conhecimento através da experiência, o qual é armazenado (tal como o cérebro), no peso das ligações entre os nós e porque têm muita elasticidade e grande adaptabilidade. Para além disso, podemos ainda acrescentar que as têm um comportamento altamente não linear, e por isso, permitem uma adaptação muito maior a informação proveniente de fenómenos não lineares.

Mas o mais fascinante, é que as RNA's são sistemas que têm um grande nível de tolerância a falhas, isto é, podem sofrer um dano considerável e continuar a ter um bom comportamento, tal como acontece nos sistemas biológicos.

Bibliografia aconselhada sobre o tema

- Cajal, S. R. 1990. *New Ideas on the Structure of the Nervous System in Man and Vertebrates*. MIT Press, Cambridge, MA, translation of the French edition of 1894.
- Canavarro Teixeira, M.C., Villamandos, N. C., Caridad e Ocerin, J. M. 2010a. Hedonic Methodologies in the Real Estate Valuation. *Congresso Internacional "Mathematical Methods in Engineering International Symposium"*, 23 outubro. Instituto Politécnico de Coimbra (Portugal).
- Canavarro Teixeira, M.C., Villamandos, N. C., Caridad e Ocerin, J. M. 2010b. Redes Neurais Artificiais para estimar o preço da habitação em Portugal. *International Meeting on Regional Science 2010*, 17-19 Novembro. Campus da Universidad de Badajoz (Espanha).
- McCulloch, W. S. e Pitts, W. 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, **5**:115-133.
- Rojas, R. 1996. *Neural Networks – A Systematic Introduction*. Springer-Verlag, Germany.
- Rosen, S. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, **82**(1):34-55.