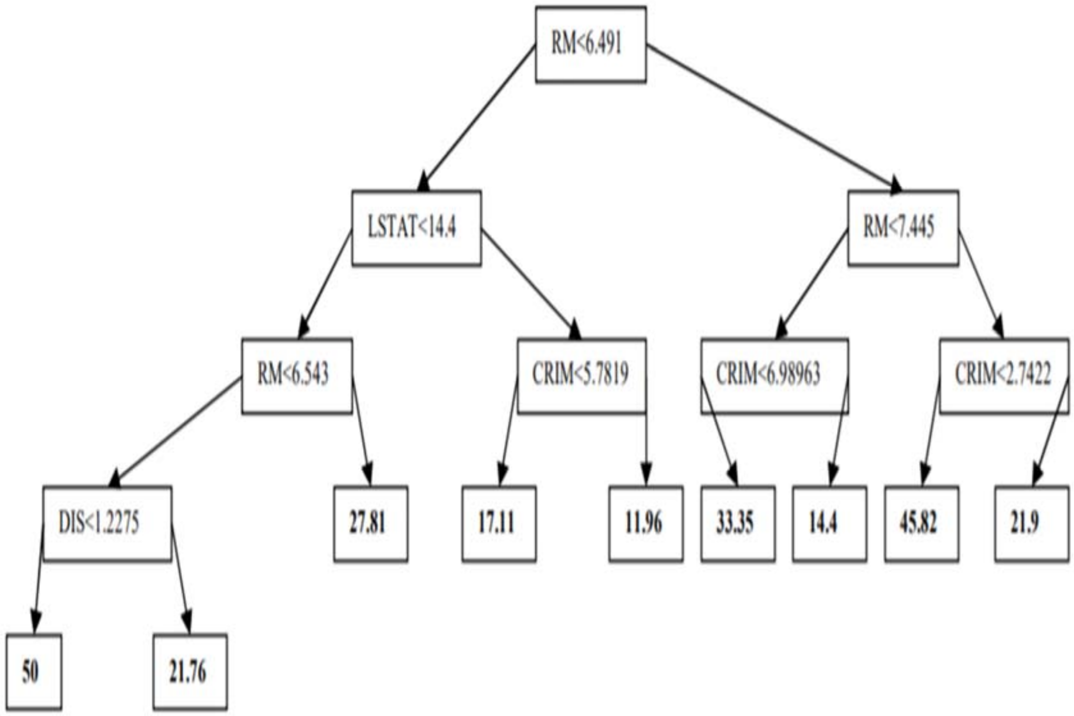
Hedonic pricing models attempt to model a relationship between object attributes and the object's price. Traditional hedonic pricing models are often parametric models that suffer from misspecification.Hedonic pricing models can also be put to fruitful use in conjoint analysis, one of the most widely applied quantitative marketing methods.The aim of conjoint analysis is to predict the utility of a new product. Similar to hedonic pricing, a product is viewed as a bundle of attributes and a mapping between these attributes and the utility represented by the product is constructed.This mapping is called the part-worth function in conjoint analysis.

Hedonic pricing hypothesizes that each good is can be looked upon as a bundle of attributes x and that a functional relationship $p = F*(x)$ exists between these attributes and the price p of a good.

Considering the relative importance plots, it is clear that there is a large difference in attribute preference across the models. The most striking differences concern the variables RAD and CRIM. RAD, the accessibility of highways, is the 4-th most important attribute according to SLR, but in the boosted CART model it occurs in the 11-th position with hardly any influence. Furthermore, boosted CART assigns a much greater weight to the attribute CRIM, the criminality ratio, than SLR. This attribute takes the 3-rd place in boosted CART and the 9-th place in SLR. The tree based models seem to rely on fewer variables than the SLR model. CART uses only 4 attributes in its model and these are also the most important ones in boosted CART. In contrast, SLR uses 11 attributes, with 7 significant ones. The partial dependence plots in Figure 7 clearly show that CART and boosted CART find nonlinearities in the data: the most important price effect of the number of rooms is between 7 and 8 rooms, and there seems to be a saturation effect in the negative price effect of LSTAT and CRIM.

Attribute	Coefficient	Std. Error	P-value
Intercept	36.341145	5.067492	$2.73 \times 10{-}12$
CRIM	-0.108413	0.032779	0.001010
ZN	0.045845	0.045845	0.000754
CHAS	2.718716	0.854240	0.001551
NOX	-17.376023	3.535243	$1.21 \times 10{-}6$
RM	3.801579	0.406316	$< 2 \times 10{-}16$
DIS	-1.492711	0.185731	$6.84 \times 10{-}15$
RAD	0.299608	0.063402	$3.00 \times 10{-}6$
TAX	-0.011778	0.003372	0.000521
PTRATIO	-0.946525	0.129066	$9.24 \times 10{-}13$
B	0.009291	0.002674	0.000557
LSTAT	-0.522553	0.047424	$< 2 \times 10{-}16$



Decision trees are usually built in two phases. The first phase is a growing phase, the second phase is a pruning phase. In the growing phase, the tree is grown until error reduction on the training set is no longer possible or a predetermined threshold has been reached. The resulting model usually overfits the data, and this is countered in a pruning phase, where the tree is shrunk until the error on a hold-out sample, the pruning set, is minimal.