

一、问题二

这里内容省略

1.1 分析该竞争关系的影响因素及其重要性

为研究我国新能源汽车与传统燃油汽车市场竞争关系受那些因素影响，我们选取新能源汽车与传统燃油车保有量的比值作为该竞争关系的量化，然后收集到 2013 年到 2022 年与新能源汽车和传统燃油车有关的因素的数据，并且使用随机森林回归算法来进行因素的重要性分析。

1.1.1 随机森林回归算法做特征变量的重要性评价的基本原理

设原始数据中有 n 个保有量比值数据、 m 个影响因素，应用自助法 (bootstrap) 重抽样技术从 n 个保有量比值数据中有放回地随机抽取 n_{tree} 个样本容量为 n 的自助样本集，从而构建 n_{tree} 株分类树，并且在每株分类树的每个节点上随机抽取 $m_{try}(m_{try} \leq m)$ 个影响因素，选择其中一个最具有分类能力的变量进行分支，且每株树都最大限度地生长，不需做任何剪枝。将所生成的 n_{tree} 株分类树组成随机森林，并将这 n_{tree} 株树分类结果的众数作为随机森林的分类结果。每次 bootstrap 重抽样未被抽到的样本组成 n_{tree} 个袋外数据 (OOB)，作为随机森林测试样本 [?]

随机森林算法对特征变量的重要性进行评价的基本思想是：对于变量 X_i ，首先，计算每株树 t 相应的袋外数据 OOB_t 的袋外误差率 $errOOB_t$ ；然后，对袋外数据中的变量 X_i 值进行随机的序列改变，而其他所有变量在保持不变的情况下，重新计算改变顺序后的袋外数据 \widetilde{OOB}_t^i 的袋外误差率 $err\widetilde{OOB}_t^i$ ，通过分析袋外数据序列改变时袋外误差的增加情况来估计某一特征变量的重要程度。变量 X_i 的重要性得分为：

$$V(x_i) = \frac{1}{n_{tree}} \sum_t (\widetilde{OOB}_t^i - errOOB_t) \quad (1)$$

1.1.2 应用随机森林回归算法对竞争关系的影响因素进行分析

为了更深入了解新能源车与传统燃油车市场竞争的影响因素，我们收集了 2013 年到 2022 年期间我国燃油价格、电动车充电成本、电动车平均价格、燃油车平均价格、新能源汽车能源效率、燃油车能源效率、新能源车政府补贴金额、新能源汽车企业数量、新能源汽车专利申请数量、新能源汽车充电桩数量以及两类车保有量之比的变化数据。将新能源汽车与燃油车保有量之比作为目标变量，将其他数据作为影响因素，运用随机森林回归模型获得每个因素的重要程度。

得到最终的重要性指标图表如下：

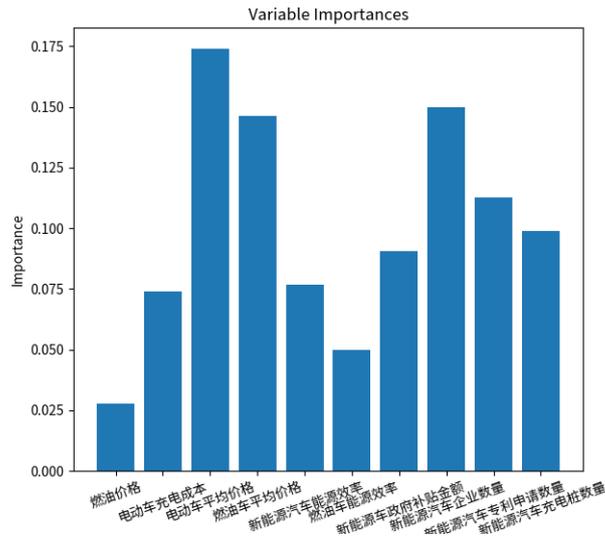


图 1: 各因素对新能源汽车和燃油车竞争关系的重要程度

在市场经济中，价格是影响产品竞争的重要因素之一。对于新能源汽车和传统燃油车这两个市场而言，价格更实惠的一方通常会在市场上占据优势。因此，我们可以看到从图??中得出的结论与之相同，即电动车平均价格和燃油车平均价格对于两者竞争关系的影响最大。

除此之外，新能源汽车专利申请数量、新能源汽车充电桩数量和电动车充电成本对竞争关系也产生了显著影响。在现代汽车技术中，专利是企业获取商业利益的重要手段之一。新能源汽车的技术含量较高，因此，专利申请数量对于新能源汽车市场的开发和扩张具有重要意义。同时，因为新能源汽车大多采用充电桩进行供能，充电桩的数量对于新能源汽车的推广至关重要。在我国新能源汽车市场起步阶段，充电桩数量是限制其推广的重要因素。由于充电桩的限制，大多数新能源汽车往往无法进行长途运行。然而，随着国家政策和新能源汽车技术的发展，充电桩问题得到了大大缓解，同时新能源汽车的能源效率也得到了显著提高，反过来缓解了新能源汽车不能长途行驶的问题。

此外，由于新能源汽车市场处于起步阶段，新能源汽车企业的集中度较低，面临的竞争压力也较小。不过，随着大量的新能源汽车企业进入市场，市场竞争日益激烈，新能源汽车在汽车市场的占有率也不断增加。

除了以上因素，燃油价格变动也能够对两者竞争关系产生影响。由于各种原因产生的全球能源危机，导致传统燃油价格上涨，这直接影响了传统燃油车的市场。同时，新能源汽车的充电成本下降，令新能源汽车的需求也随之增加。

利用此模型对测试集数据进行预测，得到如下结果：

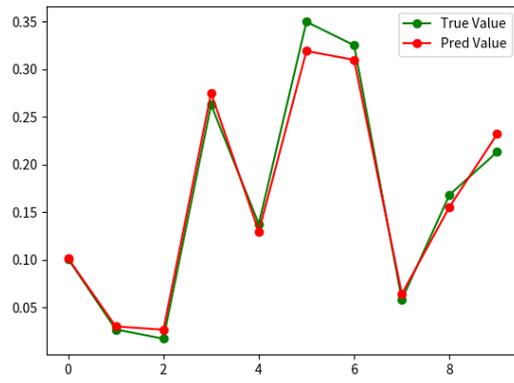


图 2: 随机森林回归模型对数据集的预测结果