图示

AI 生成的内容可能不正确。

**Full Decision Tree Interpretation**

**1. Root Split: Opening Price (Most Important Factor)**

* **Split:** open price in USD <= 3.675

**2. Left Subtree (Low Opening Price)**

* **I**f close price in USD <= 3.675:
  + Check open price in USD <= 0.85:
    - ≤ 0.85 → ~66.7% classified as competitive
    - > 0.85 → mostly non-competitive, ~76.4% classified as non-competitive
* **If** close price in USD > 3.675:
  + All competitive (0 non-competitive vs 380 competitive)

**3. Right Subtree (High Opening Price)**

* **If** close price in USD <= 10.0:
  + Mostly non-competitive, ~81.5% classified as non-competitive
* **If** close price in USD > 10.0:
  + Further split:
    - Open price in USD <= 10.068: mostly competitive, ~84.7% competitive
    - Open price in USD > 10.068: mixed, but several branches remain mostly non-competitive

**Obvious Findings**

1. Low opening price drives competitiveness. Auctions that start cheap attract more bidders.
2. High opening price discourages competition. Most high-start auctions end up noncompetitive.

**Unexpected Findings**

1. Seller reputation matters only in special cases. Among auctions with both high opening and closing prices, lower seller ratings (≤ 562) were actually more likely to be competitive.
2. Duration, end day, currency, and category barely show up in the splits.
3. Not all low-start auctions are competitive. Some low open/close auctions still ended non-competitive, possibly for undesirable items/categories.
4. High close price isn’t always competitive. At high opening prices, many auctions remain noncompetitive even if the close price is high, suggesting one serious buyer rather than multiple bidders.

**Which variable to drop. Is this model practical for predicting the outcome of a new auction?**

We drop **close price** and keep **opening price, duration, seller rating, currency, and category** as predictors. The reason is that sellers can’t control the closing price, so including it does not provide any actionable insights on how to make their auctions more competitive. Furthermore, splitting on close price is computationally burdensome. Therefore, it is reasonable and practical to exclude it from the model.

**For practical tree:** 图表, 雷达图

AI 生成的内容可能不正确。

**Practical Decision Tree Interpretation**

**1. Root Split: Opening Price (Most Important Factor)**

* **Split:** open price in USD <= 3.675

**2. Left Subtree (Low Opening Price)**

* **I**f duration <= 6.0:
  + Mostly competitive, with ~87.9% classified as competitive
* **If** duration > 6.0:
  + Check seller rating:
    - ≤ 2365.5 → ~74.3% classified as competitive
    - > 2365.5 → Mixed, if open price > 1.995, ~63.2% classified as non-competitive, otherwise, 52% competitive)

**3. Right Subtree (High Opening Price)**

* **If** seller rating <= 562.0:
  + Mostly competitive, ~63.1% classified as competitive
* **If** seller rating > 562.0:
  + Mostly non-competitive, ~69.3% classified as non-competitive

**Plot the resulting tree on a scatter plot:**

图形用户界面, 应用程序, 表格

AI 生成的内容可能不正确。图表, 散点图

AI 生成的内容可能不正确。

Splitting seems reasonable with respect to the meaning of the two predictors.

* Open Price (vertical lines): Most thresholds sit at small dollar “notches” (≈ $1–$15). Those are exactly where buyer sensitivity is steep and where sellers commonly choose starting prices. Economically, low starts spur early bids and momentum; once the opening price passes $10–$15, demand thins.
* Seller Rating (horizontal lines): Several cut points lie around a few hundred to a few thousand feedback points (≈ 500, 1k, 1.5k, 2–3k). That reflects plausible trust thresholds: credibility helps the auction attract at least two bids, especially when the opening price isn’t too high.
* Interaction structure: Price acts as the primary gate; within a price band, seller Rating refines the decision—exactly the mechanism the tree implies.

These splits are partially effective: they capture the overall trend but not cleanly.

Clear regions:

* Open Price ≤ ~$5 with Seller Rating ≥ 1k shows a noticeably higher share of competitive (orange) auctions.
* Open Price ≥ ~$15 is mostly non-competitive (blue), even with decent ratings.
* In the $5–$10 band, Seller Rating matters more: higher ratings tilt the mix toward orange.

Remaining overlap:

* At very low prices, you still see a mix of orange and blue—reputation, category, end day, and duration still influence outcomes.
* Axis-aligned splits can’t express a smooth trade-off like “if price inches up, you need proportionally more rating,” so some diagonal structure escapes the tree.

**Provide and interpret the classification table.表格

AI 生成的内容可能不正确。**

**Interpretation:**

* Accuracy = 0.701; the tree gets about 70% of auctions right on the hold-out set.
* Class 1 (Competitive)
  + Precision 0.724: when the model predicts “competitive,” it’s right ~72% of the time.
  + Recall 0.722: it captures ~72% of the truly competitive auctions.
  + F1 0.723: balanced precision/recall.
* Class 0 (Non-competitive)
  + Precision 0.674, Recall 0.676, F1 0.675: slightly worse than class 1, but close.

**Conclusion**

* The model is **practical for predicting new auctions** using only features sellers can control (opening price, duration, rating, currency, category).
* Sellers can use it to adjust auction settings to increase the chance of competitiveness, but they should be aware that predictions are not perfect.

**Based on this last tree, what can you conclude from these data about the chances of an auction obtaining at least two bids and its relationship to the auction settings set by the seller (duration, opening price, ending day, currency)?**

The following table shows the importance of predictors in the decision tree model.

Table: Predictor Importance in the Decision Tree

|  |  |
| --- | --- |
| Predictor | Importance |
| Open Price | 0.618 |
| Seller Rating | 0.297 |
| Duration | 0.080 |
| Category Music/Movie/Game | 0.005 |
| Category Books | 0.000 |
| Category Business/Industrial | 0.000 |
| Category Clothing/Accessories | 0.000 |
| Category Coins/Stamps | 0.000 |
| Category Collectibles | 0.000 |
| Category Computer | 0.000 |

The table shows that **Opening Price** is by far the most important predictor of auction competitiveness, accounting for about **62%** of the model’s decision power. **Seller Rating** is the second most influential variable (~**30%**), indicating that reputation also plays a meaningful role. **Duration** contributes modestly (**8%**), while product categories have **minimal or no impact** on the tree’s classification decisions.

Conclusion:

* **Opening price:** A lower opening price strongly increases the likelihood of a competitive auction, as it encourages more initial bidding activity.
* **Seller reputation:** Higher seller ratings are associated with greater competitiveness, holding price constant.

**Auction settings:**

* **Duration:** Five-day auctions showed slightly higher competitiveness compared to other durations.
* **Category:** Some categories, such as music, movieand **game,** appeared more competitive; however, their small sample sizes need cautious interpretation.
* **Based on the exploratory data analysis, auctions ending on Monday or Thursday tended to be more competitive than those ending on weekends, though this may vary by product category. Listings in GBP or EUR also showed slightly higher competitiveness, possibly reflecting differences in market segments and geographic factors. These insights should be applied with caution, as the sample sizes across categories vary a lot, and the decision tree did not identify ending day or currency as key predictors of competitiveness.**

**What would you recommend for a seller as the strategy that will most likely lead to a competitive auction?**

* **Set a lower opening price to attract early bidders and generate competitive momentum.**
* **Leverage seller reputation: maintain strong ratings and highlight credibility. New sellers should consider strategies to build positive feedback early.**
* **Optimize auction duration and timing: 5–7-day listings tend to perform well. 8-10-day auctions show slightly lower competitiveness. Avoid low activity ending times and consider ending auctions on Monday or Thursday.**
* **Account for category difference: adjust pricing strategies according to category-specific demand patterns and competitor pricing behavior.**