Метрика качества на тестовом наборе данных

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Алгоритм** | **Задача** | **Бейзлайн** | **Улучшенный бейзлайн** | **Самостоятельная имплементация алгоритма** |
| KNN | классификация | Accuracy: 58.60% Precision: 55.87% Recall: 58.60% F1-score: 57.17% | Accuracy: 68.22% Precision: 65.98% Recall: 68.22% F1-score: 66.54% | Accuracy: 61.81% Precision: 60.82% Recall: 61.81% F1-score: 61.23% |
| регрессия | MSE: 25690352.93 MAE: 3045.71 RMSE: 5068.57 R²: 0.83 | MSE: 22485114.74 MAE: 2595.94 RMSE: 4741.85 R²: 0.86 | MSE: 22485114.74 MAE: 2595.94 RMSE: 4741.85 R²: 0.86 |
| Линейные модели | классификация | Accuracy: 75.80% Precision: 79.35% Recall: 76.44% F1-score: 77.87% | Accuracy: 79.62% Precision: 80.72% Recall: 75.28% F1-score: 77.91% | Accuracy: 79.62% Precision: 81.10% Recall: 74.72% F1-score: 77.78% |
| регрессия | MAE: 4181.19  MSE: 33596915.85  RMSE: 5796.28  R²: 0.78 | MAE: 2756.90, MSE: 20919716.75, RMSE: 4573.81, R²: 0.87 | MAE: 3171.01  MSE: 23388209.00  RMSE: 4836.14  R²: 0.85 |
| Решающее дерево | классификация | Accuracy: 62.88% Precision: 62.08% Recall: 62.88% F1-score: 62.45% | Accuracy: 62.01% Precision: 60.48% Recall: 62.01% F1-score: 61.04% | Accuracy: 63.32% Precision: 62.02% Recall: 63.32% F1-score: 62.63% |
| регрессия | MSE: 42446908.01 MAE: 3195.11 RMSE: 6515.13 R²: 0.73 | MSE: 22812669.85 MAE: 2865.64 RMSE: 4776.26 R²: 0.85 | MSE: 20772019.06 MAE: 2688.40 RMSE: 4557.63 R²: 0.87 |
| Случайный лес | классификация | Accuracy: 80.35% Precision: 80.36% Recall: 80.35% F1-score: 80.28% | Accuracy: 81.22% Precision: 81.24% Recall: 81.22% F1-score: 81.16% | Accuracy: 78.60% Precision: 78.62% Recall: 78.60% F1-score: 78.61% |
| регрессия | MSE: 21072440.16 MAE: 2545.28 RMSE: 4590.47 R²: 0.86 | MSE: 19099851.46 MAE: 2438.07 RMSE: 4370.34 R²: 0.88 | MSE: 32235305.67 MAE: 4054.85 RMSE: 5677.61 R²: 0.79 |
| Градиентный бустинг | классификация | Accuracy: 68.56% Precision: 68.14% Recall: 68.56% F1-score: 68.15% | Accuracy: 69.43% Precision: 68.24% Recall: 69.43% F1-score: 68.76% | Accuracy: 78.17% Precision: 80.80% Recall: 79.53% F1-score: 80.16% |
| регрессия | MSE: 18944595.16 MAE: 2447.95 RMSE: 4352.54 R²: 0.88 | MSE: 18410744.19 MAE: 2422.62 RMSE: 4290.77 R²: 0.88 | MSE: 18578939.03 MAE: 2460.65 RMSE: 4310.33 R²: 0.88 |

**1. K-Nearest Neighbors (KNN)**

**Классификация:**

* **Бейзлайн:**
  + Accuracy: 58.60%
  + Precision: 55.87%
  + Recall: 58.60%
  + F1-score: 57.17%
* **Улучшенный бейзлайн:**
  + Accuracy: 68.22%
  + Precision: 65.98%
  + Recall: 68.22%
  + F1-score: 66.54%
* **Самостоятельная имплементация:**
  + Accuracy: 61.81%
  + Precision: 60.82%
  + Recall: 61.81%
  + F1-score: 61.23%

**Вывод:** Улучшенный бейзлайн значительно превосходит как бейзлайн, так и самостоятельную реализацию KNN по всем метрикам классификации.

**Регрессия:**

* **Бейзлайн:**
  + MSE: 25,690,352.93
  + MAE: 3,045.71
  + RMSE: 5,068.57
  + R²: 0.83
* **Улучшенный бейзлайн:**
  + MSE: 22,485,114.74
  + MAE: 2,595.94
  + RMSE: 4,741.85
  + R²: 0.86
* **Самостоятельная имплементация:**
  + Такие же показатели, как и улучшенный бейзлайн.

**Вывод:** Улучшенный бейзлайн и самостоятельная реализация показывают одинаково лучшие результаты по сравнению с бейзлайном.

**2. Линейные модели**

**Классификация:**

* **Бейзлайн:**
  + Accuracy: 75.80%
  + Precision: 79.35%
  + Recall: 76.44%
  + F1-score: 77.87%
* **Улучшенный бейзлайн:**
  + Accuracy: 79.62%
  + Precision: 80.72%
  + Recall: 75.28%
  + F1-score: 77.91%
* **Самостоятельная имплементация:**
  + Accuracy: 79.62%
  + Precision: 81.10%
  + Recall: 74.72%
  + F1-score: 77.78%

**Вывод:** Улучшенный бейзлайн и самостоятельная реализация линейных моделей показывают схожие улучшения по сравнению с бейзлайном, с незначительным перевесом в точности и прецизионности для самостоятельной реализации.

**Регрессия:**

* **Бейзлайн:**
  + MAE: 4,181.19
  + MSE: 33,596,915.85
  + RMSE: 5,796.28
  + R²: 0.78
* **Улучшенный бейзлайн:**
  + MAE: 2,756.90
  + MSE: 20,919,716.75
  + RMSE: 4,573.81
  + R²: 0.87
* **Самостоятельная имплементация:**
  + MAE: 3,171.01
  + MSE: 23,388,209.00
  + RMSE: 4,836.14
  + R²: 0.85

**Вывод:** Улучшенный бейзлайн значительно улучшает показатели регрессии по сравнению с бейзлайном и самостоятельной реализацией линейных моделей.

**3. Решающее дерево**

**Классификация:**

* **Бейзлайн:**
  + Accuracy: 62.88%
  + Precision: 62.08%
  + Recall: 62.88%
  + F1-score: 62.45%
* **Улучшенный бейзлайн:**
  + Accuracy: 62.01%
  + Precision: 60.48%
  + Recall: 62.01%
  + F1-score: 61.04%
* **Самостоятельная имплементация:**
  + Accuracy: 63.32%
  + Precision: 62.02%
  + Recall: 63.32%
  + F1-score: 62.63%

**Вывод:** Самостоятельная реализация решающего дерева немного превосходит бейзлайн и значительно лучше, чем улучшенный бейзлайн.

**Регрессия:**

* **Бейзлайн:**
  + MSE: 42,446,908.01
  + MAE: 3,195.11
  + RMSE: 6,515.13
  + R²: 0.73
* **Улучшенный бейзлайн:**
  + MSE: 22,812,669.85
  + MAE: 2,865.64
  + RMSE: 4,776.26
  + R²: 0.85
* **Самостоятельная имплементация:**
  + MSE: 20,772,019.06
  + MAE: 2,688.40
  + RMSE: 4,557.63
  + R²: 0.87

**Вывод:** Самостоятельная реализация решающего дерева показывает лучшие результаты регрессии по сравнению с бейзлайном и улучшенным бейзлайном.

**4. Случайный лес (Random Forest)**

**Классификация:**

* **Бейзлайн:**
  + Accuracy: 80.35%
  + Precision: 80.36%
  + Recall: 80.35%
  + F1-score: 80.28%
* **Улучшенный бейзлайн:**
  + Accuracy: 81.22%
  + Precision: 81.24%
  + Recall: 81.22%
  + F1-score: 81.16%
* **Самостоятельная имплементация:**
  + Accuracy: 78.60%
  + Precision: 78.62%
  + Recall: 78.60%
  + F1-score: 78.61%

**Вывод:** Улучшенный бейзлайн немного улучшается accuracy, precision, recall, and F1-score по сравнению с бейзлайном, тогда как самостоятельная реализация показывает снижение показателей.

**Регрессия:**

* **Бейзлайн:**
  + MSE: 21,072,440.16
  + MAE: 2,545.28
  + RMSE: 4,590.47
  + R²: 0.86
* **Улучшенный бейзлайн:**
  + MSE: 19,099,851.46
  + MAE: 2,438.07
  + RMSE: 4,370.34
  + R²: 0.88
* **Самостоятельная имплементация:**
  + MSE: 32,235,305.67
  + MAE: 4,054.85
  + RMSE: 5,677.61
  + R²: 0.79

**Вывод:** Улучшенный бейзлайн значительно превосходит как бейзлайн, так и самостоятельную реализацию случайного леса по метрикам регрессии.

**5. Градиентный бустинг (Gradient Boosting)**

**Классификация:**

* **Бейзлайн:**
  + Accuracy: 68.56%
  + Precision: 68.14%
  + Recall: 68.56%
  + F1-score: 68.15%
* **Улучшенный бейзлайн:**
  + Accuracy: 69.43%
  + Precision: 68.24%
  + Recall: 69.43%
  + F1-score: 68.76%
* **Самостоятельная имплементация:**
  + Accuracy: 78.17%
  + Precision: 80.80%
  + Recall: 79.53%
  + F1-score: 80.16%

**Вывод:** Самостоятельная реализация градиентного бустинга значительно превосходит как бейзлайн, так и улучшенный бейзлайн по всем метрикам классификации.

**Регрессия:**

* **Бейзлайн:**
  + MSE: 18,944,595.16
  + MAE: 2,447.95
  + RMSE: 4,352.54
  + R²: 0.88
* **Улучшенный бейзлайн:**
  + MSE: 18,410,744.19
  + MAE: 2,422.62
  + RMSE: 4,290.77
  + R²: 0.88
* **Самостоятельная имплементация:**
  + MSE: 18,578,939.03
  + MAE: 2,460.65
  + RMSE: 4,310.33
  + R²: 0.88

**Вывод:** Все подходы градиентного бустинга демонстрируют схожие результаты регрессии, близкие к бейзлайну и улучшенному бейзлайну, без явного преимущества самостоятельной реализации.

**Общие выводы**

1. Улучшенный бейзлайн часто показывает лучшие результаты по сравнению с первоначальным бейзлайном, особенно в задачах регрессии для линейных моделей и случайного леса.
2. Самостоятельная реализация алгоритмов:
   * KNN и Линейные модели: В некоторых случаях самостоятельная реализация показывает улучшения, однако не всегда превосходит улучшенный бейзлайн.
   * Решающие деревья: Самостоятельная реализация показывает лучшие результаты в регрессии и незначительное улучшение в классификации.
   * Случайный лес: Самостоятельная реализация хуже по метрикам классификации и регрессии по сравнению с бейзлайном и улучшенным бейзлайном.
   * Градиентный бустинг: Самостоятельная реализация значительно превосходит другие подходы в классификации, но показывает аналогичные результаты в регрессии.
3. Градиентный бустинг в классификации показывает наилучшие результаты среди всех алгоритмов при самостоятельной реализации, что свидетельствует о его высокой эффективности при правильной настройке.
4. Случайный лес и линейные модели также демонстрируют хорошие результаты, особенно в регрессионных задачах при использовании улучшенного бейзлайна.