Качество вина

Пожертвовано 10/6/2009

Включены два набора данных, связанных с образцами красного и белого вина vinho verde с севера Португалии. Цель состоит в том, чтобы смоделировать качество вина на основе физико-химических тестов (см. [Cortez et al., 2009], http://www3.dsi.uminho.pt/pcortez/wine/).

Характеристики набора данных

Многомерный

Предметная область

Бизнес

Связанные задачи

Классификация, регрессия

Тип признака

Реальный

# Экземпляров

4898

# Признаков

11

Информация о наборе данных

Дополнительная информация

Два набора данных связаны с красными и белыми вариантами португальского вина "Vinho Verde". Для получения более подробной информации см.: http://www.vinhoverde.pt/en/ или ссылку [Cortez et al., 2009]. Из-за проблем с конфиденциальностью и логистикой доступны только физико-химические (входные данные) и сенсорные (выходные данные) переменные (например, нет данных о типах винограда, марке вина, цене продажи вина и т. д.).

Эти наборы данных можно рассматривать как задачи классификации или регрессии. Классы упорядочены и не сбалансированы (например, нормальных вин намного больше, чем отличных или плохих). Алгоритмы обнаружения выбросов можно использовать для обнаружения нескольких отличных или плохих вин. Кроме того, мы не уверены, что все входные переменные релевантны. Поэтому может быть интересно протестировать методы выбора признаков.