Есть два датасета с магнитами.

В data1 данные по отдельным магнитам:

Nm - номер магнита по порядку,

Nm\_pm - номер магнита по ведомости,

Bn - магнитная индукция на северном полюсе магнита,

Bs - магнитная индукция на северном полюсе магнита,

Bav - средняя магнитная индукция магнита,

Fb - магнитный поток магнита,

Fb\_pm - магнитный поток магнита, измеренный другим способом

В data2 данные по парам магнитов, сложенных один на другой:

ind - номер по порядку,

Nm1 - номер верхнего магнита по ведомости из data1,

Nm2 - номер нижнего магнита по ведомости из data1,

Bn - магнитная индукция на северном полюсе пары магнитов,

Bs - магнитная индукция на южном полюсе пары магнитов,

Bpair\_av - средняя магнитная индукция пары,

Fn - магнитный поток на северном полюсе пары магнитов,

Fs - магнитный поток на южном полюсе пары магнитов,

Fav - средний магнитный поток пары магнитов.

Надо придумать алгоритм, чтобы так складывать магниты, чтобы для сложенных пар предсказывать значения магнитной индукции и магнитного потока.

Алгоритм состоит из нескольких шагов:

**1. Подготовка данных**

* Объединить данные из data1 и data2 по номерам магнитов (Nm\_pm в data1 и Nm1, Nm2 в data2).
* Для каждой пары магнитов в data2 добавить характеристики верхнего (Nm1) и нижнего (Nm2) магнитов из data1.

**2. Анализ и предварительная обработка**

* Проверить данные на пропуски и выбросы. Заполнить пропуски средними или медианными значениями, если это возможно.
* Нормализовать данные, если это необходимо для улучшения работы алгоритма.

**3. Выбор признаков**

Для предсказания магнитной индукции (Bpair\_av) и магнитного потока (Fav) пары магнитов можно использовать следующие признаки:

* Средние значения магнитной индукции (Bav) и магнитного потока (Fb или Fb\_pm) для верхнего и нижнего магнитов.
* Разницу или сумму значений Bn и Bs для верхнего и нижнего магнитов.
* Дополнительные признаки, такие как произведение или отношение характеристик магнитов.

**4. Построение модели**

Для предсказания можно использовать методы машинного обучения, такие как:

* **Линейная регрессия**: если зависимость линейная.
* **Метод опорных векторов (SVR)**: если зависимость нелинейная.
* **Градиентный бустинг (XGBoost, LightGBM)**: для сложных зависимостей.
* **Нейронные сети**: если данных достаточно много и зависимости сложные.

**Пример для линейной регрессии:**

python

Copy

Download

from sklearn.linear\_model import LinearRegression

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import mean\_squared\_error

*# Пример для предсказания Bpair\_av*

X = data2[['Bav\_upper', 'Bav\_lower']] *# Признаки: средние значения для верхнего и нижнего магнитов*

y = data2['Bpair\_av'] *# Целевая переменная*

*# Разделение на обучающую и тестовую выборки*

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

*# Обучение модели*

model = LinearRegression()

model.fit(X\_train, y\_train)

*# Предсказание и оценка*

y\_pred = model.predict(X\_test)

mse = mean\_squared\_error(y\_test, y\_pred)

print(f"Mean Squared Error: {mse}")

**5. Оптимизация модели**

* Подбор гиперпараметров с помощью GridSearchCV или RandomizedSearchCV.
* Использование кросс-валидации для оценки устойчивости модели.

**6. Предсказание для новых пар**

После обучения модели можно предсказывать Bpair\_av и Fav для новых пар магнитов, подставляя их характеристики в модель.

**7. Валидация и интерпретация**

* Проверить модель на тестовых данных.
* Проанализировать важность признаков (например, с помощью feature\_importances\_ в XGBoost).

**Дополнительные улучшения**

* Если данные позволяют, можно построить отдельные модели для Bpair\_av и Fav.
* Учесть физические законы, связывающие магнитные свойства пар магнитов (например, аддитивность или нелинейные эффекты).

Этот алгоритм можно дорабатывать в зависимости от конкретных требований и особенностей данных.