

Тестовое задание: Импутация SNP данных в STR

Описание задачи: В животноводстве для селекции важно иметь достоверные данные о родословной животных. Хозяйства стремятся улучшить эти данные, заказывая у лабораторий подтверждение родства между животными и их родителями. Для этого используются два типа данных: STR (Short Tandem Repeats) и SNP (Single Nucleotide Polymorphisms).

Иногда возникает ситуация, когда у потомка доступны только SNP данные, а у родителя — только STR данные (или наоборот). В таком случае для подтверждения родства требуется импутировать (восстановить) SNP данные в STR. Ваша задача — создать модель или использовать существующие методы для выполнения этой импутации по данным коров Голштинской породы.

Задача:

- Реализовать модель или алгоритм для импутации SNP данных в STR.
- Вы можете как обучить свою модель, так и воспользоваться существующими библиотеками для импутации.
- Важным аспектом является не только качество самой импутации, но и качество написанного вами кода.

Данные: По ссылке, приложенной к письму, вы найдете архив с тремя файлами:

- **FinalReport.csv** — SNP данные животных.
- **STR_train.csv** — STR данные животных, которые можно использовать для обучения модели.
- **STR_test.csv** — STR данные животных, для которых нужно выполнить импутацию.

Что от вас требуется:

1. Заполненный файл **STR_test.csv** с результатами импутации.
2. Python-код, который выполняет импутацию и позволяет получить проимпутированный файл **STR_test.csv**.
3. Python-код, который отвечает за создание и обучение модели (либо подготовку данных для использования готовых решений).
4. Краткое описание ключевых шагов, выполненных для решения задачи (в любом удобном формате).

Критерии оценки:

- **Качество импутации:** насколько точно модель восстановила STR данные.
- **Качество кода импутации:** насколько код понятен, структурирован и легко читается.
- **Код обучения модели:** логика построения модели и выполнение всех необходимых этапов (менее строгие требования к структуре кода).

Ваше решение должно быть практичным и демонстрировать ваше умение решать задачи и писать чистый код.