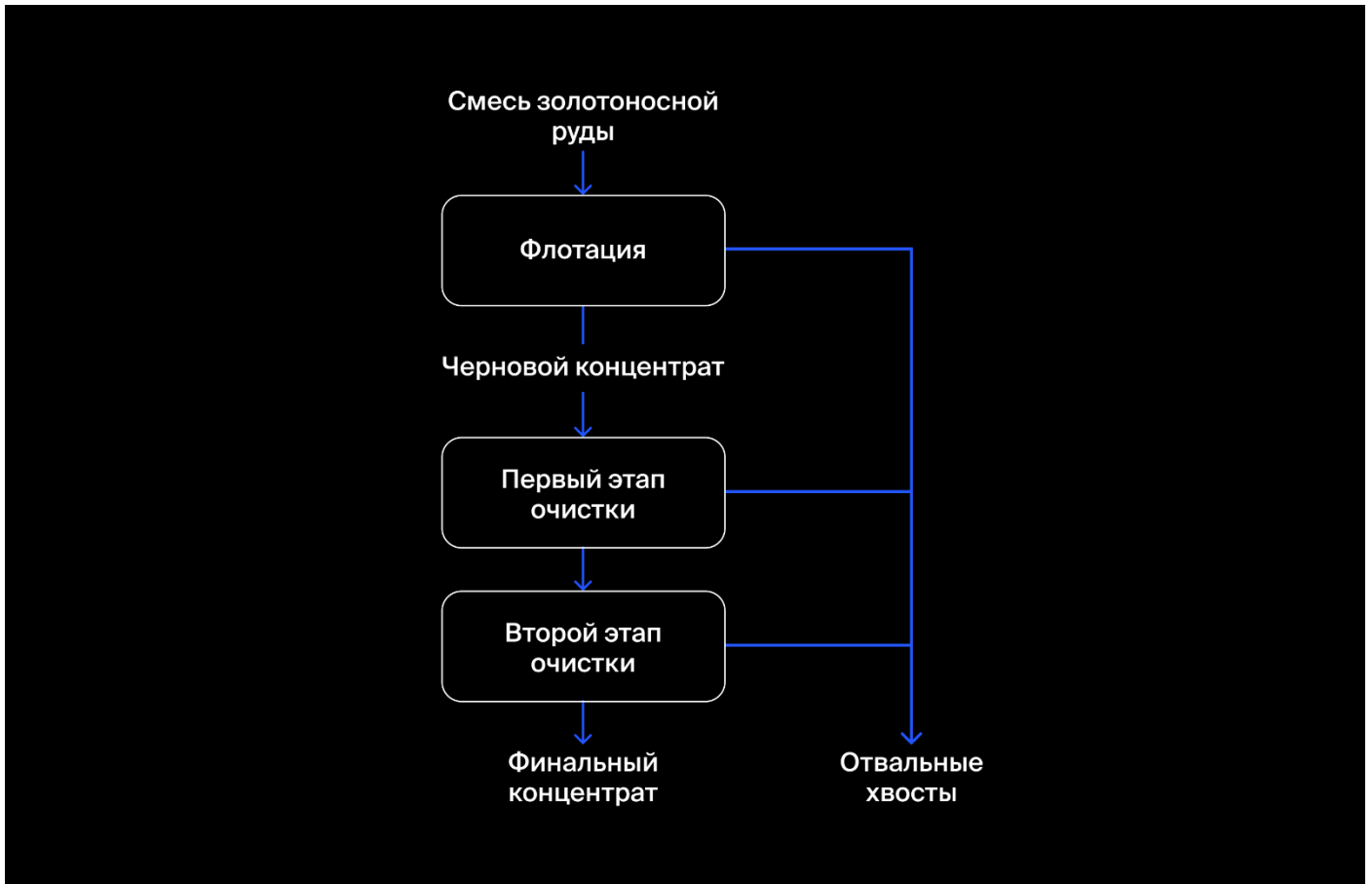


## Технологический процесс

Как золото получают из руды? Изучите внимательно стадии процесса.

Когда добытая руда проходит первичную обработку, получается дроблёная смесь. Её отправляют на флотацию (обогащение) и двухэтапную очистку.



Опишем каждую стадию:

### Флотация

Во флотационную установку подаётся смесь золотосодержащей руды. После обогащения получается черновой концентрат и «отвальные хвосты», то есть остатки продукта с низкой концентрацией ценных металлов.

На стабильность этого процесса влияет непостоянное и неоптимальное физико-химическое состояние флотационной пульпы (смеси твёрдых частиц и жидкости).

### Очистка

Черновой концентрат проходит две очистки. На выходе получается финальный концентрат и новые отвальные хвосты.

## Расчёт эффективности

Вам нужно смоделировать процесс восстановления золота из золотосодержащей руды.

Эффективность обогащения рассчитывается по формуле

$$\text{Recovery} = \frac{C \times (F - T)}{F \times (C - T)} \times 100\%$$

где:

- $C$  — доля золота в концентрате после флотации/очистки;
- $F$  — доля золота в сырье/концентрате до флотации/очистки;
- $T$  — доля золота в отвальных хвостах после флотации/очистки.

Для прогноза коэффициента нужно найти долю золота в концентратах и хвостах. Причём важен не только финальный продукт, но и черновой концентрат.

## Метрика качества

Для решения задачи введём новую метрику качества — *sMAPE* (англ. *Symmetric Mean Absolute Percentage Error*, «симметричное среднее абсолютное процентное отклонение»).

Она похожа на *MAE*, но выражается не в абсолютных величинах, а в относительных. Почему симметричная? Она одинаково учитывает масштаб и целевого признака, и предсказания.

Метрика *sMAPE* вычисляется так:

$$sMAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|) / 2} \times 100\%$$

Обозначения:

$y_i$  — значение целевого признака для объекта с порядковым номером  $i$  в выборке, на которой измеряется качество.

$\hat{y}_i$  — значение предсказания для объекта с порядковым номером  $i$ , например, в тестовой выборке.

$N$  — количество объектов в выборке.

$\sum_{i=1}^N$  — суммирование по всем объектам выборки ( $i$  меняется от 1 до  $N$ ).

Нужно спрогнозировать сразу две величины:

- эффективность обогащения чернового концентрата `rougher.output.recovery`;
- эффективность обогащения финального концентрата `final.output.recovery`.

Итоговая метрика складывается из двух величин:

$$\text{Итоговое sMAPE} = 25\% \times \text{sMAPE(rougher)} + 75\% \times \text{sMAPE(final)}$$