

Task 2

Data and privacy

Point out what data are considered sensitive personal data and advise the client on how to proceed with the planned machine learning algorithm

Here is the background information on your task

You are developing a machine-learning algorithm to predict how many days a given person will stay in the hospital. In order to do this, the hospital provides you with historical data of previous hospital visits: personal information of the patients and the lengths of their stays.

The hospital's privacy policy states that the patients' personal data can be used for developing their services, but sensitive personal data can only be used for treating the patient.

Here is your task

Before starting the development of the machine-learning algorithm, you need to decide what data to process in the algorithm.

To complete this task, answer the multiple-choice quiz. Start the quiz by clicking 'Click here to start the task' in section 5 below. The below resource will help you to tackle this task.

Start your task

You may only credit this task to your CV if you make a genuine attempt at the work.

Practice or learning attempt:



Q 2/3: You notice that health-related data sounds like a good indicator for how long a person will stay in the hospital. What do you do?

Ask the client to anonymize the data

Use the data for training the model but make sure nobody from the development team reads the data, since they are not allowed to access it.

Tell the client that this prediction model cannot be developed with their data privacy policy.

Discard all sensitive personal data and train the model on non-sensitive data such as the persons' age and location.



Q 3/3: In order to continue with the project, the hospital wants to migrate their data to a cloud-based storage. Previously, the data has been stored locally at the hospital's machines. What do you have to take in account when planning migration?

The cloud storage provider's physical data center needs to be in the EU.

The cloud storage provider must comply with GDPR.

It is not legal to store any personal data in cloud.

No special precautions are needed; you can store the data anywhere.