

BD in Healthcare Apps, Homework #4

- Answer homework questions below
- When ready, submit your answers online as instructed

1 Answer the questions

Question 1:

One can always accurately predict patient wait time from the...

1. Wait time of the previous patient
2. Current time
3. Patient age
4. None of the above

Question 2:

One of the most basic ways to build predictive models is known as

1. Linear regression
2. Linear segmentation
3. Health regression
4. None of the above

Question 3:

Adding more predictors to the predictive model always guarantees significantly better prediction quality

1. True
2. False

Question 4:

Big Data becomes particularly important in predicting clinical outcomes because

1. It is big
2. It can capture individual patient characteristics
3. It is acquired in real time
4. It is using HL7 and other clinical standards

2 Solve data problems with Matlab

As usual, let's do some big number crunching – exploring predictive models for real clinical datasets. This time, it's all about patient lines and waiting – the number one patient satisfaction criterion. In the following questions, use the attached dataset to build different predictive models, and to estimate their quality.

Keep in mind that linear regression always uses the constant intercept term (constant predictors), in addition to any other predictors you select. Make sure you have the intercept in all your analyses.

Useful hints:

- Use Matlab `regress(Y,X)` function to predict Y from X
- In our case, Y is the column vector of observed patient waits. X contains columns of predictor values
- The first column in X should be a column of ones (the intercept) – let's call it I. Then, if you need to predict from the current line size (vector L0), you will have to use

```
[b,bint,r,rint,stats] = regress(Y, [I L0]);
```


This will return the vector r of residual errors (prediction errors), and this is the vector you will need to use to find the accuracy of your model.
- stMRN in your data file is a unique patient identifier (Medical Record Number).

Question 5:

Model 1: Predict patient wait time from the intercept alone. For this model, all patients are given the same predicted wait time. No additional predictors are used.

What percentage patients have their wait time predicted with 5-minute accuracy by this model?

Enter your result... (for example, enter 30 for 30%)

Question 6:

Model 2: Now expand the model, adding current line size in number of patients. What percentage patients have their wait time predicted with 5-minute accuracy?

Enter your result... (for example, enter 30 for 30%)

Question 7:

Model 3: Now expand the above model, adding current line size and line size 5 minutes before the patient arrived. What percentage patients have their wait time predicted with 5-minute accuracy?

Enter your result... (for example, enter 30 for 30%)