Rules:

You NEED access to a computer with Internet connection and Python installed (+NumPy).

You are allowed to use all course materials and all static material in the Internet. You are NOT allowed to communicate with anyone in any form during the exam.

You MUST join the course Slack channel during the exam and you can send private messages to the course instructor.

All code and text MUST be written by yourself. Copying substantial parts from any source is PLAGIARISM.

Online test has multiple deadlines. If you miss a deadline, provide wrong material or incorrect answers, then the test stops there.

DO NOT re-submit files for the previous deadlines as it resets their time tags and your test is over.

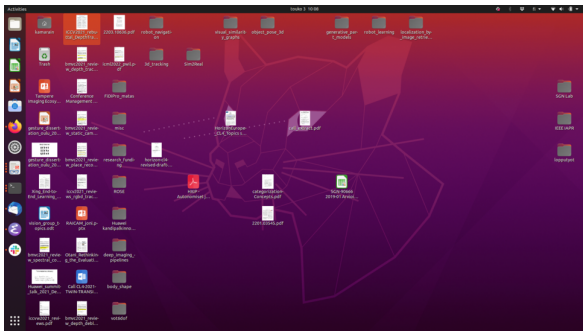
1. Basics (1pts) **[Return before 9am10]**

   Take two screenshots of your desktop:

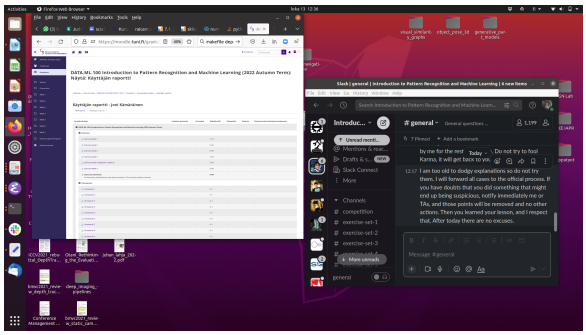
   - An empty desktop where all applications are minimized (see Fig. 1(a))
   - Open your grading page in Moodle DATA.ML.200 (name visible) and the course Slack (see Fig. 1(b))

   Submit the following items before the deadline:

   - Screenshot 1 as PNG file: `surname_desktop1.png`
   - Screenshot 2 as PNG file: `surname_desktop2.png`



(a) Empty desktop



(b) Moodle and Slack open

Figure 1: Examples of the screenshots that should be returned

2. Regression **[Multiple deadlines]**

   Download the training and test data files (`X_train.dat` and `X_test.dat`), and their target values (`y_train.dat` and `y_test.dat`). You can use numpy.loadtxt() function to load the data files.

   (a) Baseline (2pts) **[Return before 9am30]**
   The regression task is to estimate the continuous output y_i from 14-dimensional input $\vec{x}_i$, i.e. $y_i = f(\vec{x}_i)$
   As the baseline, we use the mean value of the training outputs y_i.
   Write Python code to calculate mean absolute error for test data and print it:

```
(dataml100) Joni$ python kamarainen_baseline.py
Reading data... Done!
Computing baseline regression.
  Baseline accuracy (MAE): x.xxx
(dataml100) Joni$
```

Submitted items:

- Screenshot of the full desktop that shows the results of running your code: `surname_baseline_screenshot.png` (Figure 2)
- Python code: `surname_baseline.py`

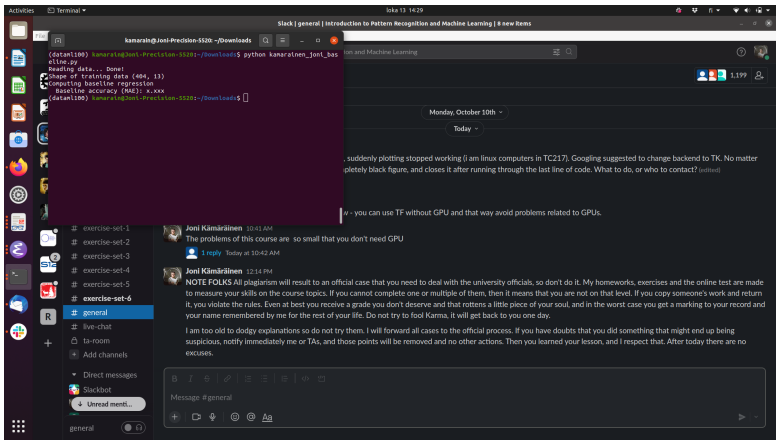


Figure 2: Example screenshot

(b) 1 nearest neighbor regression (2pts) **[Return before 10am00]**
The next regressor, f_{1nn}, is based on the nearest neighbor rule.

Write Python code that for each test sample searchs for its closest training sample using Euclidean distance, and then gives the closest sample y as the estimate $\hat{y}$.

Again, compute MAE over all test samples:

```
(dataml200) Joni$ python kamarainen_1nn_regression.py
Reading data... Done!
Shape of training data (404, 13)
Computing 1-nn  regression
  Baseline accuracy  (MAE): x.xxx
  1NN regr. accuracy (MAE): y.yyy
(dataml200) Joni$
```

Submitted items:

- Screenshot of the desktop with a window (terminal) running the code: `surname_1nn_regression_screenshot.png`
- Python code: `surname_1nn_regression.py`