Exercise 4

1. The historical student enrolments at the Alabama University are given in the file **alabama.txt**, containing the year and number of students enrolled that year. Fit a linear regression model with least squares method on this data, predicting number of enrolled students (y) with the year (x). Plot the data and regression line in same plot. What is the predicted number of students for year 2050 according to this model? Looking at example images in slides on Page 10-11 in **Prediction.pdf**, would linear regression be a good choice for modelling these time-series? Do not use existing implementations of linear regression for this task.
2. Using the same **alabama.txt** data, design a fuzzy prediction system using (simplified) Song's method. For further details see Song and Chissom "Forecasting enrollments with fuzzy time series - Part I/II", *Fuzzy Sets and Systems 54*, 1993.
    - Create equally spaced, triangular fuzzy sets over range of number of enrolments (**Prediction.pdf** Page 17)
    - Fuzzify historical data, where each year's enrolments belong to fuzzy set with highest membership values.
    - Create prediction rules (Page 20). If there are multiple consequents, combine them with "OR" ($\cup$) operation.
    - For each year in historical data, predict the next year's enrolments. Use centroid-of-area for defuzzification.
    - Plot these predictions and original data alongside each other. Is the prediction good?
    - Try different number of fuzzy sets in original step to find good prediction results. What happens if number of these sets is too low or too high?