Self Evaluation

02460 Advanced Machine Learning

Supervisor: Morten Mørup
Group members:
Daniel Horvath: s172185
Lorant Gulyas: s172188
Laura Perge: s182257

Kernel Density Estimation with Expectation Maximization Applied on Randomly Missing Value Imputation

Predefined Learning Objectives:

1) Recall Kernel Density Estimation (KDE), identify its use cases and limitations, highlight that its usability is overshadowed by the curse of dimensionality.
2) Define the version of the Expectation Maximization (EM) algorithm that extends the KDE's density estimation in a way that it is less prone to the previously established issue. Outline and exemplify potential constraints and solutions.
3) Code implementation of the version of the EM algorithm derived in point 2) for KDE, and integrate both hold-out and leave-one-out cross-validation within the iteration loops.
4) Test the code to measure and detect causes of inefficiency and integrate solutions where possible. Increase computational performance by preparing an alteration of the classical Multivariate Gaussian PDF based on the Cholesky decomposition.
5) Derive and illustrate how the variational inference (VI) can be integrated into the KDE. Contrast it with the EM algorithm, emphasize and discuss the ability of the VI to return approximate distribution of the covariance instead of a point estimate.
6) Hypothesize that the optimized version of the KDE enhanced with the EM algorithm can be used to impute randomly missing values in case of higher dimensionality. Measure and assess the performance of the enhanced KDE by experimenting on datasets.
7) If time allows, debate the losses coming from restricting the covariance matrix to be strictly diagonal.
8) Build a compact but comprehensive report about our findings.

Evaluation:

1) Completed
   See Section 1, Introduction
2) Completed
   See Section 2.2
3) Completed
   See GitHub, EM_KDE, not optimized
4) Completed
   See GitHub , EM_KDE, optimized
5) Incomplete
   Not enough time and no need for the application case (See Week 10-11)
6) Completed
   See Sections 2.3, 2.4, 3, 4; GitHub Imputation
7) Completed
   See Sections 2.3, 2.4, 3, 4
8) Completed
   See Report