

Popular science summary of the PhD thesis

PhD student Niels Bruun Ipsen

Title of the PhD thesis Generative Modelling with Missing Data

PhD school/Department DTU Compute

Science summary

* Please give a short popular summary in Danish or English (approximately half a page) suited for the publication of the title, main content, results and innovations of the PhD thesis also including prospective utilizations hereof. The summary should be written for the general public interested in science and technology:

Missing values forces the analyst to make choices, either explicitly or implicitly, about how to proceed with a given analysis. These choices are mostly perceived to be of practical nature, but often tacetly imply analytical assumptions. The challenges that missing data imposes can be handled in a statistically principled manner by marginalizing over the missing data in probabilistic models.

We develop and analyze probabilistic generative models with latent variables in missing data problems:

- In a simple analytically tractable case the effect of missing values on parameter estimates is analysed in detail
- When the mechanism leading to missing values depends on the missing values themselves, the missing mechanism needs to be modelled. We introduce a modelling approach utilizing the tools of amortized variational inference to model the observed data and the missing mask jointly.
- In supervised learning no distribution over the covariates is typically assumed and the usual approach of marginalizing over the missing features is not possible. We investigate different methods for handling missing values in supervised deep learning and propose an approach to marginalizing over missing features in a joint model of targets and covariates

Please email the summary to the PhD secretary at the department