COL333 Assignment 4

Raja Kumar 2021CS10915
Aastha Rajani 2021CS10093

November 29, 2023

Underlying Concept

The implementation focuses on probabilistic inference in a Bayesian network using the Expectation-Maximization (EM) algorithm. The approach assumes equal probabilities for all entries in the conditional probability table and iteratively performs expectation and maximization steps.

EM Algorithm

- **Expectation (E) Step:** For each data point, the algorithm calculates the expected values of latent variables (hidden variables) given the observed data and the current parameter estimates. These expected values are then used to estimate the likelihood of various configurations of latent variables.
- **Maximization (M) Step:** The parameters (conditional probability tables) of the Bayesian network are updated based on the expected values obtained in the E-step. The primary goal of the algorithm is to maximize the likelihood of the observed data given the current parameters.

Convergence and Time Allocation

- The algorithm iterates through E and M steps until a convergence criterion is met or until the allocated time is exhausted. Convergence indicates that the parameters have stabilized, suggesting that further iterations are unlikely to significantly improve the model.

Additional Insights

- To handle edge cases, the implementation assumes a minimum probability of 0.01. Through multiple random trials, the best results consistently occurred with a smoothing factor of 0.0035.