Yuha Yi

Stat 517

Due 11/30/18

Literature Critique for Taylor Dean: Support Vector Machines with Feature Selection in Breast Cancer Diagnosis

This literature will be split between the presentation and the paper. Taylor Dean’s presentation was clear and understandable from a perspective that does not include a biological studies background. He beings with a meaningful introduction in why the topic is relevant and it’s importance followed by the usage of predictive modeling in the topic. The visualizations used were clear and explained well in explaining the support vector machines. The explanation of the mathematical and statistical background for the explanation of the support vector machine’s functions were lack luster and only provide a brief explanation in the presentation. However, given the time setting this is appropriate, and the mathematics are explained in the paper which will be reviewed below. The features are shown and explained clearly. Taylor explains why the author chooses the features by evaluating them using the F statistic while providing the equation for this. The slide that includes the model parameters was vague to me. Perhaps an explanation of the kernel selection would have been useful. The performance evaluation was very thorough by explaining the different types of evaluation that included equations for sensitivity, specificity, positive predictive rate, negative predictive rate, and roc curves. The explanation for each was clear and the equations were easy to follow, along with a table to explain the positives and negatives. Model selection was very straightforward to the results. In the conclusion of the results, Taylor included the scoring along with the area under the curve which was great. Overall performance of the presentation would be a 9 out of 10.

The language used in the paper is very formal for a research setting. Once again there is a quick and strong introduction with a flawless implementation of the topic and how the statistics is involved. The mathematics is explained in a broad sense for an overview of the support vector machines. My suggestion here would be to explain the coefficients in the equation. As an example, explain the function phi(w), how it relates to the margin, and why the margin distance is important. It is difficult to follow the implementation of the equation into the research. However, the F-score is explained in depth along with how it is integrated in the paper. Once again as in the presentation the evaluation methods is explained well along with the choosing of the evaluation and why the final decision was made. The summary of the conclusion brings back all the main points from the paper and Taylor includes the future work that the paper had. In a macro sense the overall paper is structurally well built with clear organization. There was only a small flaw in the explanation of some of the mathematical equations.

The strongest point in Taylor’s literature review was clarity of the literature. The weakest point was the comprehensiveness of some of the mathematical equations. The presentation and paper were very well done.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Length of Paper | Depth of Understanding | Completeness | Difficulty | Comprehensiveness | Clarity |

10/10 9/10 10/10 10/10 8/10 10/10