Final Project Presentation Review of G6

Review submitted by G10

Title: Spatial Decision Tree With Application In Land Cover Classification

Problem Motivation:
The project falls under applied research and the objective of the project was to improve the existing land cover classification approaches by introducing the concept of spatial decision trees. They mention in their motivation slide that land cover classification can be used for pipeline planning, environment science etc. However, the motivation behind the problem was not explained clearly and was not elaborate. The slide on motivation moreover focusses on topics which are more relevant to be placed in the challenges section (auto-correlation, noise etc.) and the novelty section (traditional decision trees ignore spatial relationships). There are a number of applications of spatial land cover classification, such as determining the extent of urbanization around a city or a town, finding the crop patterns and other agricultural trends in a farm, monitoring and expansion of forest types and shrubland etc. The motivation behind the problem could have touched upon some of these aspects. The topics listed in the motivation slide are being repeated in other sections as well which could have been avoided.

Problem Statement
The objective of the problem was not clearly understood from the problem statement slide as spatial decision tree model was not known beforehand but became more apparent in the slides that followed on first principle example. The inputs, outputs, objective and constraints could have been more explicitly mentioned. Feature maps and class maps were not familiar terms and it could have helped if the slide included their definitions. A distinction can be kept between challenges and constraints. For example, spatial auto-correlation, noise and large size of the data can be called challenges which the approach should be able to handle. Constraints are the limitations of the model under which the model is expected to perform. Assumptions are the simplifications induced in the model which helps in designing the approach and improving upon its computational efficiency. For example, training samples are representative of test samples is an assumption. The first principle example was very clear and intuitive in explaining the problem statement.

Challenges
The challenges were articulated clearly and the computational challenges in developing a spatial decision tree model was shown by example figures/numbers. However, it would have been better if spatial entropy would have been defined and discussed in detail at this stage as it is being employed in the proposed approach as well. The challenges in using noisy feature maps can be shown with the help of some example feature map figures.

Proposed Approach
The slide on the approach included the pseudo code of the algorithm which was a little over technical and difficult to understand. It would have helped if they could have explained the approach in simpler terms and given an algorithm trace over a toy dataset for clarification. The neighborhood auto-correlation based spatial entropy was not very clear and could have been explained in detail. In the approach, a split is performed in the spatial decision tree favoring the maximum information gain using the spatial entropy measure. The approach appears impressive but could be communicated better.
Novelty
Related work was not complete as it did not include traditional decision trees in the classification framework which was being compared with their approach in the first principle examples. Related work focused more on the classification of techniques which employed spatial entropy – distance based, joint-probability based and spatial neighborhood based. However, an explanation of how spatial entropy was different from traditional entropy and the context of its use in each of three spatial entropy based approaches could have been given.

Validation
Experimental results on a given ground truth observation of land cover classification was used to compare the performance of their approach with traditional decision trees. However, as pointed out in the discussions that followed in the end of the presentation, the evaluation metric used was ROC curve which does not take into account the spatial auto-correlation in the data which computing the precision and recall. Hence, the evaluation could be improved by using spatial accuracy measures of performance, which can incorporate spatial auto-correlation.

Presentation Critique
The talk was useful and the approach proposed was novel. However there is a scope of improvement and the presentation can be rated 8/10.

1. Was the talk accessible to an "intelligent lay person"? 7/10. Simpler language than pseudo code could have been used to explain the approach. Also, some definitions could have been given in the beginning for better understanding, such as feature map, class map and spatial entropy. Also a running trace of the algorithm could have been given for better clarity.
2. Did the talk emphasize a central message that conveys the overall value of the work being executed? 9/10 The central message was that spatial auto-correlation should be employed for developing decision trees for land cover classification when the feature maps can be noisy and choppy.
3. Did the talk attempt to relate to the audience and showed effort in conveying key ideas clearly? 8/10. The first principle example explained the problem effectively. However, the same first principle example could have been used for explaining the running of the algorithm.
4. Was the speaker's response to questions satisfactory? 8/10 It was pointed out that precision and recall measures may not hold correct in capturing spatial auto-correlation in the data and hence better performance measures should be employed.
5. How did the talk do on covering the 6 elements? Kindly rate each element separately and include a brief justification for each.
   1. Motivation: 7/10. Could have been more elaborate as the topics presented in the motivation slide are more suited to be placed in challenges or related work.
   2. Problem Statement: 8/10. It was explained well through first principle example. However, the constraints could be improved and more introductory definitions could be provided.
   3. Challenges: 8/10. Explained through numerical examples. However, some figures could have helped in explaining the challenges better.
   4. Proposed Approach: 8/10. The approach is well-structured. However, simpler language could have helped express the approach better. Also spatial entropy could have been defined in detail.
   5. Novelty: 8/10. Could have included traditional decision trees in the classification
   6. Validation: 8/10. Precision recall measures can be replaced by their spatial counterparts.