Project Proposal Presentation Review (Csci 8715: Spatial Data Science)

Date	3/20/26
Presenter group	G3 – Woochang Shin, Yashasvi Singh
Reviewer group	G2 – Noah Constable, Joe Conroy

Sr.	Review Questions	Constructive Feedback
1.	**Provide a brief summary of the proposal listing the problem** (e.g., definition, significance, challenges) **and the proposed approach** (e.g., description, novelty, superiority over competition)**.**	**This proposal identifies the problem of detecting food deserts, that is areas lacking access to affordable, nutritious foods (i.e. grocery stores as opposed to fast food). They state that previous attempts to solve this problem use buffer models and fail to capture deeper spatial patterns, such as the interaction between different food outlet types. They call this a 'structural' imbalance. The authors propose a hybrid approach, combining spatial co-location mining with hotspot detection. The novelty lies in this approach and the identification of anti-co-location that can be analyzed with hotspot detection to create meaningful categories like "fast-food dominant" regions.**
2.	**Is the problem stated clearly?** - Does it list inputs, output, an objective function, and constraints (e.g., key assumptions). - Does it illustrate inputs and output. - Does it define and illustrate key concepts needed to understand the problem statement?	**The authors state the problem well. The inputs are defined as geolocated food outlet data (spatial), a neighborhood definition, and an interaction radius constant R. The outputs are defined as classified neighborhoods and co-location metrics. Constraints are identified as using continuous spatial data, sensitivity to R definition, and an expected uneven urban density. This is mostly listed in the slides, but the paper could be improved by explicitly defining the objective function, as right now it doesn't seem formally defined. Illustrations help in the slideshow, but the paper could use one end-to-end example.**
3.	**Is the problem's importance articulated by addressing questions such as the following:** Who cares about the problem? If the project is successful, what difference (e.g., societal, technical) will it make?	**The proposal clearly identifies who cares about the problem (city officials, public health officials, civilians) and ties the problem to societal outcomes like inequity and diet-related diseases. This 'why' part is one of the stronger parts of the paper!**
4.	**Are the problem's key challenges identified?**	**Challenges are identified: defining neighborhoods in continuous space, sensitivity to R, data sparsity, capturing absence (anti-co-location).**
5.	**Is an approach proposed to address the key challenges in the problem?**	**The proposed hybrid approach directly addresses the stated challenges and shows a good understanding of the material.**
6.	**Is the novelty of the proposed approach articulated?** For example, did the proposal summarize related work and their limitations overcome by the proposed approach?	**The proposal clearly states that related work has not focused on anti-co-location or this approach, emphasizing novelty, but could be stronger with a direct example of where previous methods fail ("approach A fails in scenario X, while our approach doesn't because…").**
7.	**Did the proposal articulate the superiority of the proposed approach over the state of the art?** For example, did it provide an example (or other preliminary results or other arguments) comparing the proposed approach with related work.	**Superiority is partially supported by the plan to compare against a 1-mile buffer method and to analyze metrics (precision, recall, IoU). A toy example or small illustrative comparison would help support the claim of superiority.**
8.	**Did the proposal describe an evaluation plan to gather additional scientific evidence to assess the proposed approach?** Did it list evaluation questions (e.g., correctness, completeness, computational complexity, comparative analysis, sensitivity analysis, scalability, etc.)?	**The listed evaluation plan addresses:** **- correctness using synthetic data experiments to verify that metrics behave as expected** **- comparative analysis against a 1-mile buffer model and local government reports** **- sensitivity analysis using a varying neighborhood radius (300-500m)**

	- Did it name scientific methodologies (e.g., proofs, computational simulation and experiments, real-world tests (e.g., road-test), case studies)? - Are named scientific methodologies appropriate for the evaluation goals?	- **performance metrics using precision, recall, IoU, PI stability** **Some missing evaluations that could be included are computational complexity/scalability and an evaluation of robustness using noisy data. The proposal does clearly list:** - **computational simulation** - **real-world empirical testing** - **comparative benchmarking** **as scientific methodologies and these are appropriate for the task at hand.**
9.	**Did the proposal include a work plan?** - Does it decompose the proposed work into a set of (weekly) tasks? - Does it provide a schedule for the tasks? - Did it list milestones (mid-term and final "exams") to check for success? - Did it list risks (in the proposed approach)?	**Work plan is clear, structured, and realistic. Tasks are broken down weekly, in a scheduled order. It does miss some specific milestones, such as "validate PI impl. by week 5" and it is missing a discussion of risks. These additions would make the overall weekly schedule slightly more robust.**
10.	**Other comments**	**Some terminology could be better defined to help the understanding of a reader and to strengthen significance claims. Two stuck out to me. The first being the class labels posed in the Proposed Approach section – "Balanced Service, Fast-Food Dominant, Grocery Desert, or True Food Desert". In this case, "fast-food dominant" and "grocery desert" sound like they are capturing the same relationship – a high proportion of fast-food POIs to grocery POIs. The proposal should clearly define the criteria separating the classes.** **Second, the term "structural absence" is central to the proposal but not formally defined. It would help to clarify whether this strictly refers to missing food service categories within the POI set, or a broader notion of missing urban services. Providing a precise definition (possibly tied to anti-co-location metrics) would improve clarity.**