*Review*

# Spatial Decision Support Systems with Automated Machine Learning: A Review

**Richard Wen [1]* and Songnian Li [1]***

[1] Department of Civil Engineering, Toronto Metropolitan University, Toronto, Canada; rwen@ryerson.ca (R.W.); snli@ryerson.ca (S.L.)
* Correspondence: rwen@ryerson.ca (R.W.); snli@ryerson.ca (S.L.)

**Abstract:** Many spatial decision support systems suffer from user adoption issues in practice due to lack of trust, technical expertise, and resources. Automated machine learning has recently allowed non-experts to explore and apply machine learning models in the industry without requiring abundant expert knowledge and resources. This paper reviews recent literature from 136 papers, and proposes a general framework for integrating spatial decision support systems with automated machine learning to lower major user adoption barriers. Challenges of data quality, model interpretability, and practical usefulness were discussed as general considerations for system implementation. Research opportunities related to spatially explicit models in AutoML, and resource-aware, collaborative/connected, and human-centered systems were also discussed to address these challenges. This paper argues that integrating spatial decision support systems with automated machine learning can not only encourage user adoption, but also mutually benefit research in both fields — bridging human-related and technical advancements for fostering future developments in spatial decision support systems and automated machine learning.

**Keywords:** Spatial; Decision Support; Machine Learning; Automation; Framework; System; SDSS; AutoML; GIS

## 1. Introduction

Advances in crowdsourcing [1], open data initiatives [2], and open source standards [3] have made spatial data more publicly accessible. Spatial Decision Support Systems (SDSS) store, manage, and process spatial and non-spatial data for important decisions, such as selecting business locations, placing traffic infrastructure, and implementing public health policies [4]. However, many SDSS are not adopted by decision makers due to lack of trust, technical expertise, and resources [5,6]. Recently, Automated Machine Learning (AutoML) received attention from the research community and media. AutoML integrates automation and Machine Learning (ML) by generating models, with little human assistance, that perform well under certain requirements and computational budgets [7]. This reduces the effort and technical expertise required to process and model data, which accounts for a majority of the time spent on data analysis [8]. As leading technology companies released AutoML products in 2017 to 2018 [9–11], ML models became more widely used by non-experts and less expensive to implement. With the recent increase of accessibility to AutoML, resources for implementing and maintaining SDSS can be reduced — improving SDSS adoption by decision makers.

This paper provides a systematic review of AutoML and SDSS integration, which seeks to answer three research questions: (1) What problems can both SDSS and AutoML solve according to recent research? (2) How can AutoML be integrated into SDSS? and (3) What are the challenges and opportunities of SDSS with AutoML? Although there are existing review papers on AutoML and SDSS separately [4,12**?** ], review papers that focused on the integration of both AutoML and SDSS were not found in literature from initial search on the topics of AutoML and SDSS together. This paper produces the following three research contributions to answer these questions: (1) A systematic review that investigates recent methods, results, applications, and potential problems in SDSS and AutoML (2)

check for updates

**Citation:** Wen, R.; Li, S. Spatial Decision Support Systems with Automated Machine Learning: A Review. *Preprints* **2022**, *1*, 0. https://doi.org/

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A framework based on recent literature for implementing AutoML in SDSS and (3) A summary of key research opportunities and challenges of SDSS with AutoML.

## 2. Methods

This paper used a two-step process to answer three research questions described in the introduction (Figure 1). The process involved gathering relevant AutoML and SDSS research literature, then summarizing, analyzing, and discussing the gathered literature to answer the three research questions.

### 2.1. Step One: Literature Search

The first step involved a search of recent AutoML and SDSS literature. Peer reviewed journal articles were keyword searched (titles only) in 382 research databases (e.g. Scopus, arxiv, Web of Science, etc) using Summon 2.0 between January 1, 2019 to September 24, 2022. These articles were then manually filtered by inspecting the title and abstracts for AutoML/SDSS relevant and literature review related articles only. Since many recent literature review articles covered developments in AutoML [12**?** –15] and SDSS [4,16–19] within the past 5 years, articles within two years were used to avoid outdated information and to focus on the most recent research. The manually filtered articles (17 AutoML, 18 SDSS) were used to also discover 101 additional supplementary references for both AutoML and SDSS topics using a snowball search strategy [20].

### 2.2. Step Two: Review and Discussion

The second step involved answering the research questions using the articles and relevant references gathered from the first step. Similar problems solvable by AutoML and SDSS were identified and summarized to answer research question one: *What problems can both SDSS and AutoML solve according to recent research?* A general framework for SDSS with AutoML was developed to answer research question two: *How can AutoML be integrated into SDSS?* Similarities in the methodology sections from the review papers in step one, and relevant references, were examined to identify AutoML and SDSS components. The AutoML and SDSS components were then connected based on general AutoML/SDSS approaches from the review papers/relevant references, and identified spatial problems from answering research question one. Finally, opportunities and challenges from the review papers and relevant references in step one were identified and discussed to answer research question three: *What are the challenges and opportunities of SDSS with AutoML?* Challenges and opportunities were found by comparing similarities and differences among the results, discussion, limitations, and other related sections/references.

## 3. Results

A total of 136 SDSS and AutoML articles were found for review. 18 SDSS and 17 AutoML articles were used as primary sources to review recent important advancements, while 63 SDSS articles, 21 AutoML, and 17 AutoML/SDSS articles (articles having both AutoML and SDSS as topics) were found using these primary sources to serve as supplementary sources for reviewing foundational literature earlier than 2022. The number of articles per year are steady between 1990 and 2018 (supplementary selection for most notable publications), before sharply rising after 2018 (primary selection for more recent articles) (Figure 2). Articles with both AutoML and SDSS topics were quite recent and seen only after 2020. Main keywords were focused on the topics of data, spatial/planning systems, machine learning, and models/analysis (Figure 3). Notable articles based on the number of citations are seen in Figures 4 and 5. Notable primary SDSS articles had over 15 citations:

- *Spatial Decision Support Systems: Three decades on* [4]
- *Reporting on the Performance and Usability of Planning Support Systems—Towards a Common Understanding* [21]
- *Advances of Four Machine Learning Methods for Spatial Data Handling: a Review* [22]

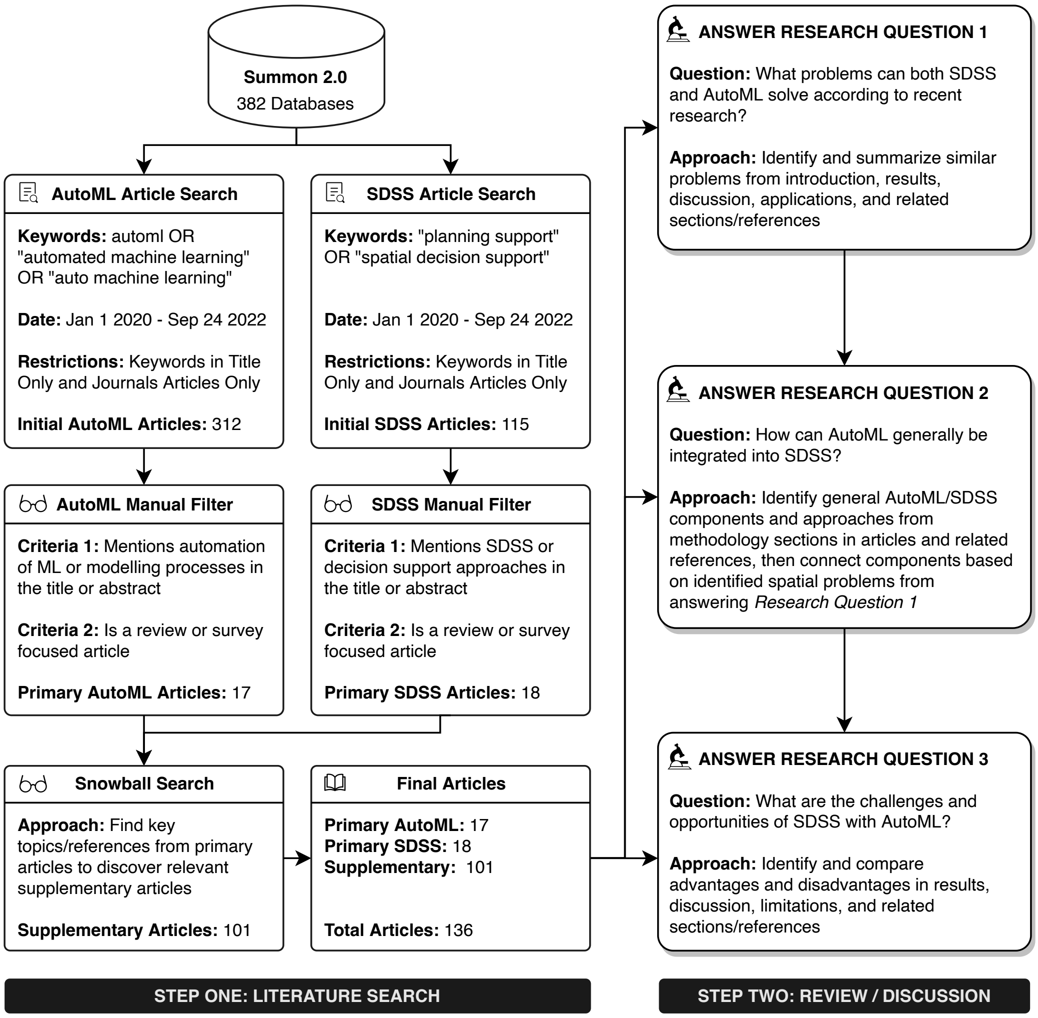


**Figure 1.** Two step process to answer research questions.

- *Strengthening Participation Using Interactive Planning Support Systems: A Systematic Review* [17]

Notable primary AutoML articles had over 25 citations:

- *Automated machine learning: Review of the state-of-the-art and opportunities for healthcare* [23]
- *AutoML: A survey of the state-of-the-art* [12]
- *Benchmark and Survey of Automated Machine Learning Frameworks* [24]

Notable supplementary SDSS articles had over 500 citations:

- *Interpretation of the Correlation Coefficient: A Basic Review* [25]
- *Kriging: a method of interpolation for geographical information systems* [26]

Notable supplementary AutoML articles had over 1000 citations:

- *Random Forests* [27]
- *What is a support vector machine?* [28]
- *K-means clustering: A half-century synthesis* [29]

Notable supplementary articles with both AutoML and SDSS topics had much less citations due to recency and had over 5 citations:

- *Reconstruction of GRACE Total Water Storage Through Automated Machine Learning* [30]
- *Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning* [31]

- *Machine Learning to Estimate Surface Roughness from Satellite Images* [32]

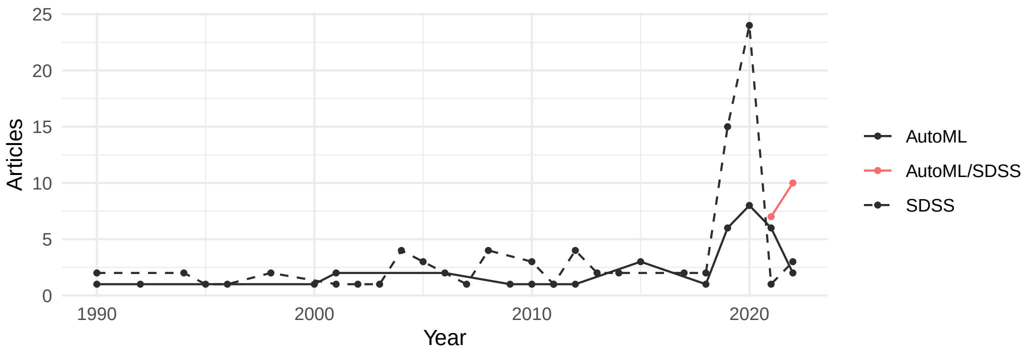


**Figure 2.** Final SDSS and AutoML Articles Per Year (n=136).

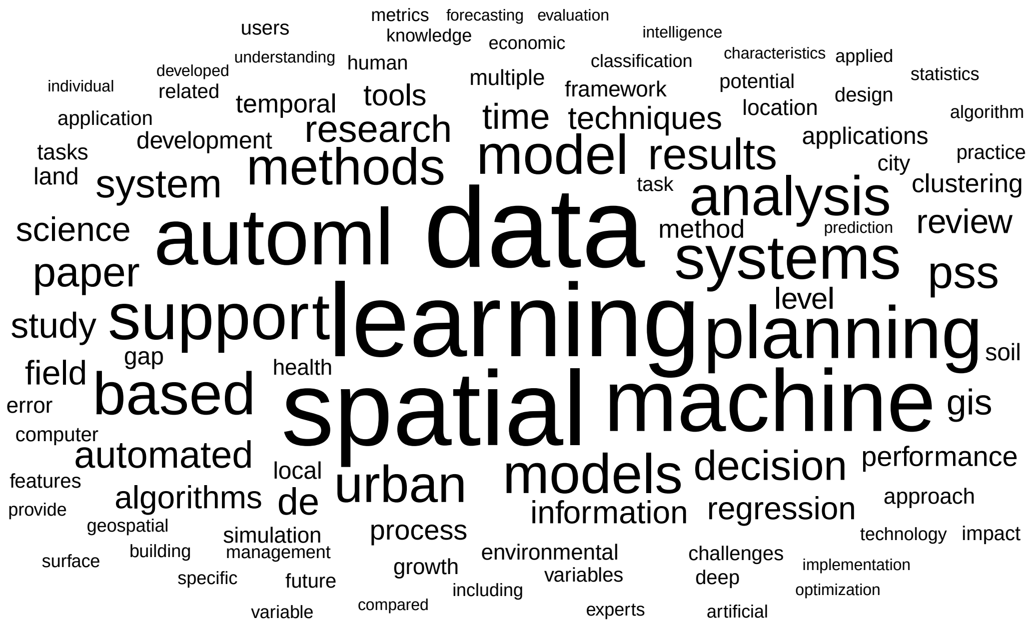


**Figure 3.** Word cloud of top 100 words in 136 SDSS and AutoML article abstract and titles.

## 4. Literature Review

### *4.1. Spatial Decision Support Systems (SDSS)*

The term SDSS was used since 1985 to describe software designed to support decision making by enabling users to analyze structured or semi-structured spatial problems for potential solutions [33]. Modern SDSS shifted from solution-centric software to human-centric frameworks that incorporated features and ideas from Geographic Information Systems (GIS) [34] and Planning Support Systems (PSS) [35]. SDSS are frameworks that incorporate a collection of tools designed to inform decision making involving spatially related problems, generally comprised of three components [4,16,17,21,36]: (1) spatial data (2) spatial information and (3) knowledge.

The spatial data component manages and processes data as input for the spatial information component, which transforms the data into information (e.g. modelling, visualization) by organizing and presenting it for decision making needs [19,37,38]. Notable approaches used in the spatial information component include Multiple Criteria Decision Analysis (MCDA) [39], hot spot analysis [40], spatial regression [41], Cellular Automata (CA) [42], Agent Based Modelling (ABM) [43], and Particle Swarm Optimization (PSO) [44].

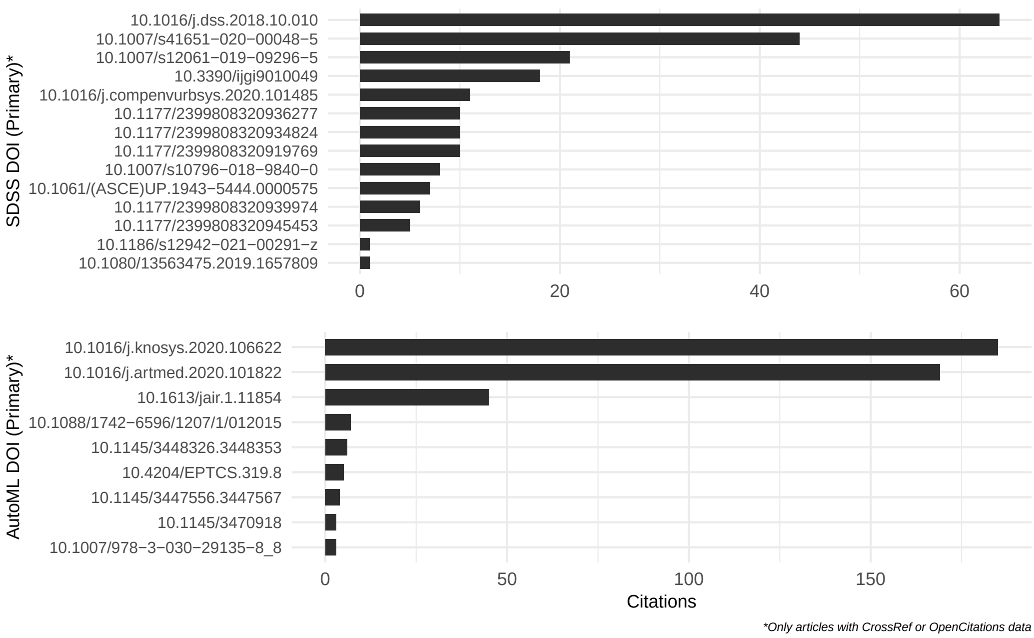


**Figure 4.** Primary AutoML (n=17) and SDSS (n=18) article citation counts.

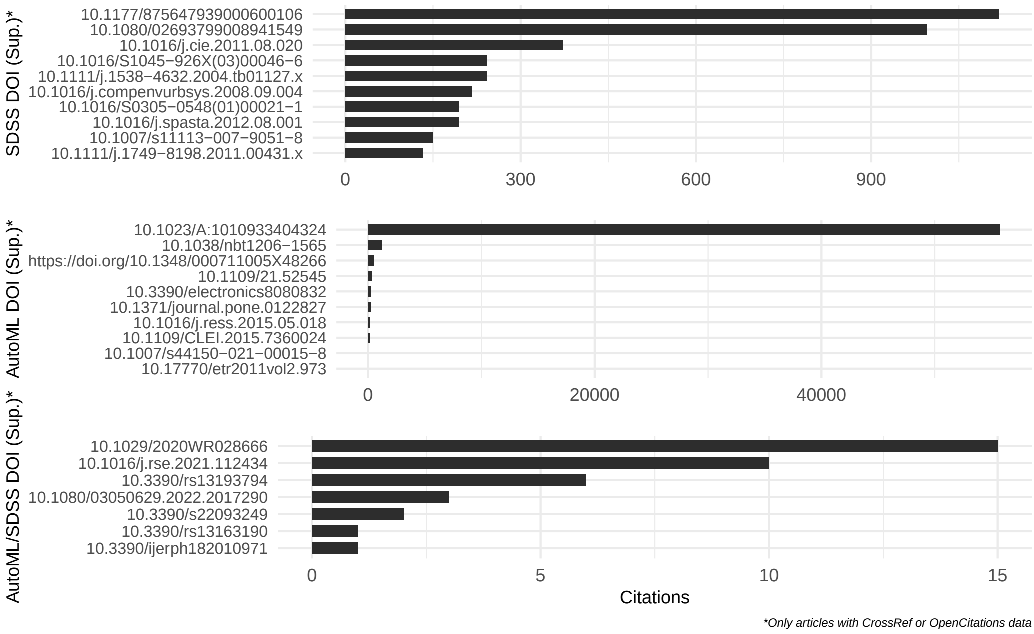


**Figure 5.** Top 10 most cited supplementary AutoML (n=21), SDSS (n=63), and AutoML/SDSS (n=17) articles.

This information acts as input for the knowledge component, which enables users to interact with and explore the information to produce knowledge [18,45–48]. Knowledge is used to support decisions or to help improve the data or information components [4,21]. Notable approaches for knowledge generation include participatory planning [17], citizen science [45], and geocollaboration [49]. Although the spatial data and information components focus on spatial data, they may also include supplementary non-spatial data. Each SDSS component contains several subcomponents representing more specific component features and functionality. General SDSS components, subcomponents, and their interactions are seen in Figure 6.

 doi:10.20944/preprints202209.0465.v1



Despite many studies in SDSS [4], challenges exist involving low user adoption [5,48,50] (e.g. awareness across disciplines, lack of practitioner acceptance/trust, expensive resources/training), evidence of usefulness [16,47] (e.g. proving practical utility/success for practitioners, added value vs resources needed), adaptation [19] (e.g. balance between domain-specificity and generalizability, application to similar contexts/problems), collaboration [36] (e.g. communication between non-technical and technical actors, translation of decision making needs to models/tools), and interpretability [6] (e.g. excessive/complex information for non-technical users, transparency of processes/inputs/outputs). Discussion around the gaps between research and practice of SDSS have remained an important topic with many studies encouraging the early involvement and collaboration between decision makers, stakeholders, and the community [5,36,51]. Over time, SDSS research, largely focused on applications, case studies, and reviews, has started to increased after 2004 and remained steady between 2010 and 2020 with a recent growth in studies related to urban science/analytics, smart cities/urban planning, and digital twins [52].

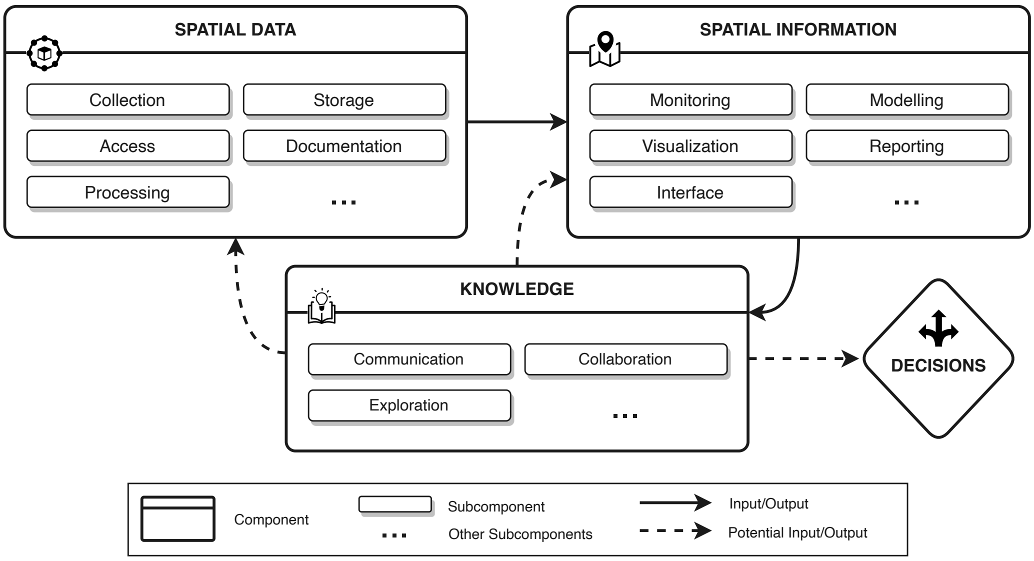


**Figure 6.** Spatial Decision Support System (SDSS) components.

### *4.2. Automated Machine Learning (AutoML)*

AutoML can be described as the automation of machine learning, a combination of two terms: (1) *automation* (Auto), to independently act, function, or operate without human intervention [53] and (2) *Machine Learning* (ML), a field of Artificial Intelligence (AI) focused on computer algorithms that can improve through experience [54]. Current AutoML approaches commonly involve optimizing components in the ML process (e.g. extraction/creation of features, tuning/creation of models) given constraints (e.g. reaching desired performances or time limits) [**?** ]. These optimization approaches use metrics (e.g. accuracy, error) that determine the quality of the output components (e.g. features, models). Notable models include linear/logistic regression [55,56], Naïve Bayes (NB) [57], Decision Trees (DT) [58], Random Forests [27], k-means clustering [29], Support Vector Machines (SVM) [28], Neural Networks (NN) [59], and Genetic Algorithms (GA) [60]. A generic AutoML approach is seen in Figure 7.

Although AutoML has made AutoML more accessible to non-experts [7,61], there are issues in data dependency [12,23] (e.g. low data quality, unavailable data, data misuse), time and efficiency [15,62] (e.g. performance vs acceptable running times, dataset sizes, search space comprehensiveness), updates/reusability [63**?** –65] (e.g. update existing model with new data, performance consistency, reproducible solutions), and interpretability [12,66,67] (e.g. why models perform better/worse or take certain actions). Closing the gap between

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1



domain experts/practitioners and ML specialists has recently been a topic of interest as ML processes are increasingly automated, and applied to solve practical problems in the industry [68]. Much of AutoML research has been focused on supervised learning, but recent research has diverged to tackle a larger range of ML problems such as unsupervised learning, time-series forecasting, and anamoly detection [69,70].

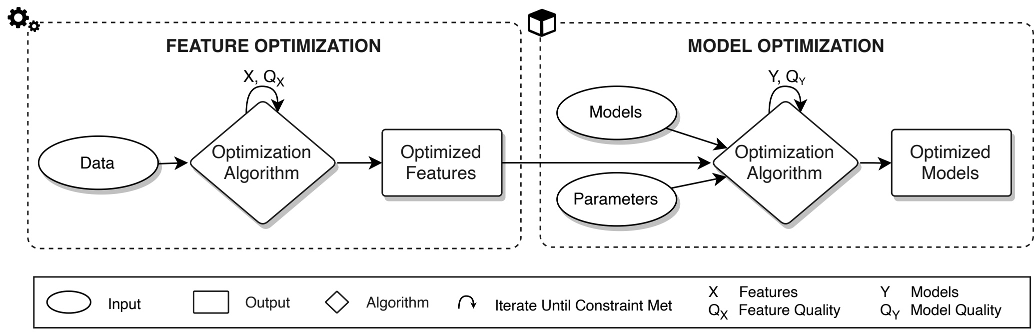


**Figure 7.** Generic Automated Machine Learning (AutoML) approach.

### *4.3. Spatial Problems in SDSS and ML*

Spatial problems, commonly studied in SDSS, were not prevalent in AutoML research before 2020, where many studies focused on general problems, such as prediction and optimization, without considering spatial effects. However, ML applications to spatial problems are more common [71,72], and are recently being integrated into SDSS [19,73]. This section identifies spatial problems that have been studied by SDSS and/or ML approaches to supplement the much smaller number of studies focused on both SDSS and AutoML. Since AutoML automates ML processes, it is relevant to review studies that use ML to solve spatial problems. A summary of spatial problems in SDSS and reviewed applications and approaches is shown in Table 1.

**Table 1.** Reviewed spatial problems and approaches in SDSS.

| **Spatial Problem** | **Applications** | **Spatial Methods** | **ML Methods** |
|---|---|---|---|
| Estimation | Land use classification<br>Disease risk calculation<br>Disaster risk prediction | MCDA<br>Spatial Regression | SVM<br>RF<br>NN |
| Optimization | Facility selection<br>Delivery routing<br>Infrastructure placement<br>Traffic control | MCDA<br>PSO | GA |
| Clustering | Crime hotspots<br>Agricult./disease zoning<br>Social media analysis<br>Travel analysis | LISA<br>Hotspot analysis<br>SaTScan | K-means<br>DBSCAN |
| Simulation | Traffic control simulation<br>Wildfire simulation<br>Landuse simulation | Cellular Automata<br>ABM<br>Custom models | RL |
| Insight | Risk factor identification<br>Interactive exploration<br>Data/model interpretation | Spat. Regression<br>Webmapping<br>Interactive models | Feat. selection<br>Feat. importance<br>Pipeline explore |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1



#### 4.3.1. Spatial Estimation

Spatial estimation problems involve the calculation of unknown values at different locations, which encompasses spatial interpolation [74], prediction [75], and overlay [76]. These problems are solved to create surfaces from samples (e.g. kriging [26]), predict future values at different locations, and calculate values based on location. Recent examples include disease risk calculation [77], disaster risk prediction [78], and land use indicator creation [79] using MCDA approaches, spatial regression (e.g. Geographically Weighted Regression (GWR) [80]), and ML (e.g. SVM, RF, NN). Spatial estimation solutions are often evaluated with error (e.g. Root Mean Square Error (RMSE), true negatives, Sum of Square Error (SSE)) and accuracy (e.g. F1 score, Area Under the ROC Curve (AUC), sensitivity, specificity) metrics that compare estimated values to true values from real world data [81]. Challenges in spatial estimation problems involve need for larger amounts of ground-truth data, and consideration of additional factors/variables/scales [75].

#### 4.3.2. Spatial Optimization

Spatial optimization involves spatial placement [82] and routing [83] of entities. Solving spatial optimization problems help determine more optimal placement of important facilities and infrastructure (e.g. site selection or facility location problem [84]), and efficient transportation paths (e.g. vehicle routing [85] and travelling salesman problems [86]). Recent examples include hospital facility selection [87], energy infrastructure placement [88], delivery routing [89], and automatic traffic control [90] using GA, MCDA, and PSO. In addition to error and accuracy metrics, spatial optimization solutions are evaluated with multicriterion [91] (e.g. weighted sums, sensitivity analysis) and multi-objective [92] (e.g. spread, hypervolume, convergence) metrics that consider important indicators from data, expert input, and search space comprehensiveness. Challenges in spatial optimization problems involve temporal effects, multiple objectives, and computing efficiency [84,93].

#### 4.3.3. Spatial Clustering

Spatial clustering involves grouping entities in space (e.g. Local Indicators of Spatial Association (LISA) [94], hot spot analysis, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [95]) and time (e.g. spatiotemporal clustering [96], SaTScan [97], K-means [29]), where entities inside groups are similar and entities outside different groups are dissimilar [98]. These groups are used to identify interesting areas (e.g. zones with high crime [99] or natural resource potential [100]) and timespans (e.g. areas of disease transmission during specific times [101]). More recent applications involve clustering high volume and velocity spatiotemporal data (e.g. social media, crowdsourcing) [102,103]. Spatial clustering solutions are evaluated with statistical [104] (e.g. likelihood ratio, autocorrelation, significance) and similarity [105] (e.g. euclidean distance, pearson correlation) metrics. Spatial clustering challenges include irregularly shaped clusters, high dimensional data, spatial relations/weights selection, resolution, object interactions, and visual vs quantitative evaluation [96,106,107].

#### 4.3.4. Spatial Simulation

Spatial simulation refers to imitations of real-world or hypothetical phenomena in space and time [108]. Spatial simulation enables analyses in cases where data is difficult to obtain (e.g. finer spatial/temporal resolutions, hypothetical/future phenomena) [43]. Approaches involve domain-specific models (e.g. crop yield models [109], hydrodynamic fluid models [110]), CA, and ABM. Recent examples include Reinforcement Learning (RL), ABM, and CA for simulating traffic light control [111], wildfire spread [112], and sustainable urban growth [113]. Spatial simulation solutions are evaluated with domain-specific (e.g. total traffic delay [111], total crop yield [109], landscape composition/ patch sizes [114]) metrics that are used to guide empirical observations. Challenges in spatial simulation include model validation, excessive complexity, disorganization, reproducibility, computing resources, and lack of theoretical basis [115].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1



#### 4.3.5. Spatial Insight

Spatial insight problems focus on interpreting and visualizing spatial data and model outputs, typically involving spatial regression [116], interactive interfaces [117], and maps [118]. Incorporating spatial insight features in SDSS help data and models produce knowledge useful for decision making (e.g. graphical interface to interactively view and manipulate spatial data and models [119], web maps displaying spatial data or model results [120], coefficients representing variable effects in models [121]). Recent examples include spatial regression for identifying factors for reducing pollution [122], web GIS to interactively generate watershed models [123], and interactive visualization tools for exploring and analyzing AutoML pipelines [124]. Spatial insight solutions are evaluated with variable-based (e.g. feature importance [125], coefficients[25]), interpretability [126] (e.g. cognitive indicators, explanation indicators), and empirical approaches [127] (e.g. usability testing, controlled user surveys, user insight observation) to examine the effectiveness of interactive/visualization tools for producing useful insight. Spatial insight challenges involve measuring usefulness, selecting/justifying appropriate presentation methods, handling big data/complexity, and personalization vs generalizability [127–129].

### *4.4. SDSS with AutoML*

Research related to SDSS with AutoML recently begun to emerge after the year 2020. AutoML methods were applied to a variety of applications in the areas of agriculture (e.g. crop prediction [130–132], crop classification [133]), environmental science (e.g. environmental impact assessment [134], waterlogging risk estimation [135], water storage estimation [30], water potential mapping [136], meteorological forecasting [137], ocean behaviour prediction [138]), geology (e.g. oil well placement [139], soil roughness estimation [32], soil moisture estimation [31], landslide risk estimation [140,141]), transportation (e.g. road health inspection [142]), and public health (e.g. violence rate prediction [143]). The majority of studies use a fusion of geospatial data sources with satellite imagery, sensors, and surveys being the most common and sociodemographic data being the least common. A summary of SDSS with AutoML approaches and applications is shown in Table 2.

The reviewed AutoML methods were grouped into four generalized approaches: (1) Ensembling, (2) Bayesian, (3) Neural Nets, and (4) Evolutionary. The most prominent AutoML approach was ensembling, which involves the combination of multiple algorithms to achieve better performance than individual algorithms [144]. This is followed by Bayesian approaches that are based on the Bayes theorem and utilizing past observations to guide future predictions [62,145]. Neural net approaches involve the optimization of neural network architecture to build deep learning models that achieve high performance [146]. Lastly, evolutionary approaches use algorithms that mimic natural selection based techniques on a population of models, such as mutation, reproduction, and selection to find optimal evolved models that achieve better performance [147].

Among the reviewed AutoML methods and software, notable ones include Auto-Sklearn, Tree-based Pipeline Optimization Tool (TPOT), H2O, Autogluon, Neural Architecture Search (NAS), and Alpha3DM. Auto-Sklearn uses a combination of Bayesian optimization, meta-learning, and ensemble construction to perform hyperparameter tuning and algorithm selection [148]. TPOT uses genetic programming to optimize and generate tree-based ML pipelines [149]. H2O uses random search and stacked ensembles to produce a final model that evaluates a diversity of candidate models, which, in some situations, are better than Bayesian optimization or genetic algorithm based approaches [150]. Autogluon distills ensembled models into individual models, by using a data augmentation strategy based on Gibss sampling, to produce final models that are faster and, in some cases, more accurate than training individual models by themselves or ensembled models [151]. NAS automates the construction of neural network architectures to build deep learning models with different search strategies, spaces and performance estimation techniques, which can be applied to Bayesian Nets and Long Short Term Memory (LSTM) networks [146]. AlphaD3M uses meta reinforcement learning with self play by modelling meta-data, tasks,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1



and ML pipelines as state in a deep learning model, which allows AlphaD3M to be faster than AutoML approaches such as Auto-Sklearn and TPOT while being explainable with transparent ML pipeline edit operations [152].

**Table 2.** Reviewed SDSS with AutoML approaches and applications (n=17).

| AutoML Approach | AutoML Method/Software | Data | SDSS Applications |
|---|---|---|---|
| Ensembling (n=6) | H2O<br>Extra trees class. | Satellite imagery<br>UAV imagery<br>Sensors<br>Surveys<br>Sociodemographic<br>Simulations | Crop prediction<br>Violence rate prediction<br>Total water storage est.<br>Landslide risk est.<br>Soil estimation |
| Bayesian (n=4) | Bayesian Opt.<br>Auto-Sklearn<br>Bayesian Nets<br>MATLAB fitrauto | Satellite imagery<br>UAV imagery<br>Sensors<br>Surveys | Crop prediction/classify<br>Soil estimation<br>Env. impact assessment |
| Neural Nets (n=3) | NAS<br>Deep Learning<br>LSTM | Satellite imagery<br>Vehicle imagery | Meteorological forecasting<br>Road health inspection |
| Evolutionary (n=2) | Fedot<br>TPOT | Sensors<br>Surveys | Oil well placement<br>Waterlogging risk est. |
| Other (n=2) | Autogluon<br>Auto$^n$ML<br>AlphaD3M | Sociodemographic<br>Satellite imagery<br>Surveys | Violence rate prediction<br>Water potential mapping |

## 5. Discussion

### *5.1. SDSS with AutoML Framework*

Recent research in SDSS has incorporated ML algorithms and other models, which can be automated by AutoML, to solve spatial problems for transforming spatial data to spatial information. Applying the generic AutoML concept in Figure 7 to the SDSS components in Figure 6, AutoML can be integrated into SDSS by framing these spatial problems as optimization problems. Given spatial problem $x$, potential solutions $S$, and metric $Q$ (measuring how well solutions solve the spatial problem), AutoML can auto-approximate near optimal solutions $\tilde{S}$ for $x$ based on metric $Q$ within pre-defined constraints (e.g. time limits, desired performance). For example, a spatial estimation problem $x$ could be to classify whether pixels at different coordinates are urban or rural land use. Metric $Q$ can be a measure of how many pixels were predicted correctly based on groundtruth samples of urban/rural land use pixels, while potential solutions $S$ can be a set of appropriate models (e.g. kriging, DT, NN) that can predict urban/rural landuse pixels. The near optimal solution $\tilde{S}$ is the most accurate model from $S$ based on $Q$, using an optimization algorithm (e.g. GA, PSO) under constraints (e.g. max iterations/runtime). In MCDA, the potential solutions $S$ can also be different weighing schemes. A framework to integrate AutoML into SDSS is seen in Figure 8, where AutoML automatically processes spatial data into spatial information by solving various spatial problems. In general, this framework requires three key considerations:

1. **Spatial Problems**: What are the spatial problem(s) to solve given the context and actors in decision making?
2. **Metrics**: What metrics are appropriate for evaluating and measuring the defined spatial problem(s)?

 doi:10.20944/preprints202209.0465.v1



3. **Potential Solutions**: With the given spatial problem(s) and metric(s), what are the potential solutions to solve the spatial problem(s)?

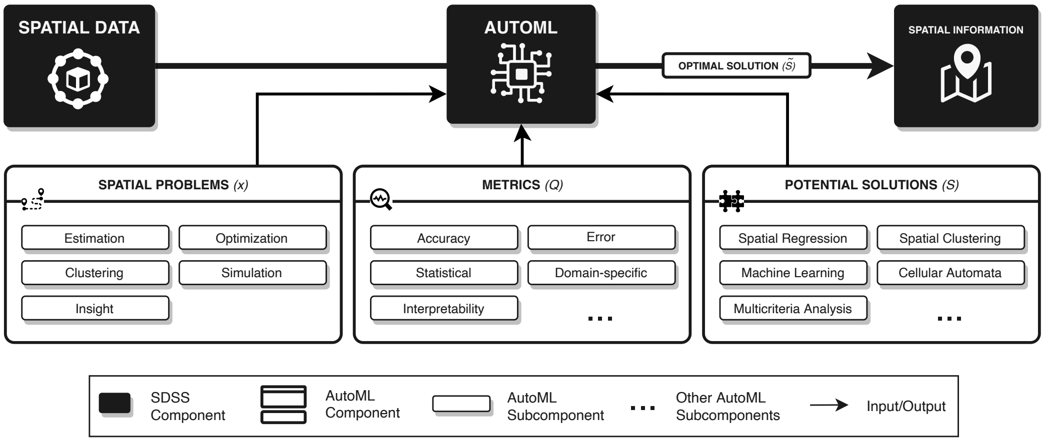


**Figure 8.** General SDSS with AutoML Framework.

5.1.1. Key Consideration 1: Spatial Problems

Given the context and actors in decision making, spatial problems need to be defined to reflect decisions being evaluated [153]. Initially, determining the types of decisions being evaluated may help in defining the behaviour of decisions. The main types of decisions according to [154] are:

- **Independent**: decisions made by a decision maker with full responsibility and authority
- **Sequential Interdependent**: decisions made partially by a decision maker and partially by another party
- **Pooled Interdependent**: decisions made from negotiation and interaction among decision makers

Then three main steps related to the process of decision making as described by [155] can further aid in defining spatial problems:

1. **Intelligence**: Examination of spatial data to identify spatial problems that require decisions and have the opportunity for change
2. **Design**: Determining possible and alternative decisions and developing approaches to evaluate and understand the decisions
3. **Choice**: Selecting from the range of possible and alternative decisions after evaluating and understanding each decision

After considering the type of decisions, the possible decisions, and approaches to selecting/evaluating the possible decisions, the spatial problem may be better defined as one or more (but not limited to) of the following general spatial problems as reviewed in Section 4.3:

- **Spatial Estimation**: calculation of unknown values in space (e.g. prediction, overlay)
- **Spatial Optimization**: optimization of entities in space (e.g. placement, routing)
- **Spatial Clustering**: organization of entities in space (e.g. grouping, categorization, zoning)
- **Spatial Simulation**: simulation of phenomena in space (e.g. physics, theoretical simulations)
- **Spatial Insight**: interpretation and exploration of phenomena and entities in space (e.g. interactive maps, visualizations, plots)

Decisions can be simple or complex depending on the change desired from the decision and the evaluation approach designed. A simple decision may only require defining a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1



single spatial problem. For example, the identification of crime hotspots for police patrols can be defined as a spatial clustering problem. A more complex single decision may require defining one or a combination of different spatial problems. For example, analyzing the effects of health interventions may require the identification of intervention areas (a spatial clustering problem) and simulating the effects of each alternative intervention on the intervention areas (a spatial simulation problem). The considerations in this section are meant to be a starting point to help structure decisions as spatial problems, but not as a definitive guide to do so as every context, actor, and decision making process can vary drastically and there is not a solution for every situation [153].

#### 5.1.2. Key Consideration 2: Metrics

After defining spatial problems to solve, the metrics used to measure and evaluate potential solutions to each spatial problem need to be determined. It is important to consider the task needed to solve the defined spatial problems. For ML and statistics, the following tasks are common amongst studies [81,156,157]:

- **Regression**: estimation or prediction of continuous target values given other factors (e.g. calculating landslide risk, predicting number of traffic collisions)
- **Classification**: identification of discrete target values (groups or categories) given other factors (e.g. predicting landuse types, identifying building types)
- **Clustering**: organization of entities into groups or categories based on characteristics (e.g. identifying crime zones, disease areas)

Considering the tasks for solving spatial problems aids in identifying the metrics required to evaluate each spatial problem. Each metric has a particular purpose and is suited for measuring the performance of particular tasks. For example, RMSE and correlation coefficients metrics are used to evaluate regression tasks, while accuracy and F1 scores are used to evaluate classification tasks. As an example metrics for the mentioned tasks above are provided in Table 3. [81] presents a more comphrehensive collection of regression and classification metrics, while [105], [158], and [159] provide more detailed overviews of different clustering metrics.

**Table 3.** Example metrics for spatial problems and tasks.

| Task | Metrics |
|---|---|
| Regression | Error, Correlation Coefficient, MSE, MAE, RMSE |
| Classfication | Accuracy, Precision, Recall, Sensitivity, Specificity, F1 Score, AUC, ROC |
| Clustering | Euclidean Distance, Rand Index, Entropy, Purity, Silhouette Coefficient, Dunn's Index, Calinski-Harabasz Index, Homogeneity |

The choice of metrics should fit not only the defined spatial problem, but the behaviour and characteristics of the spatial data used. It is important to note that each metric has its own advantages and caveats [81,92,158]. For example, accuracy metrics are biased when the data contains unbalanced classes (e.g. 90% of the data is class A and only 10% is class B). This causes classification models to be high performing if a large majority of the output from classification contains the dominant class. In this case, the F1 Score is a more appropriate metric to account for class imbalances.

#### 5.1.3. Key Consideration 3: Potential Solutions

When the spatial problems and associated metrics are defined, potential solutions to match these problems and metrics can be determined. The spatial problems and metrics create a structured goal for AutoML approaches to reach, where the potential solutions

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1



are often possible models or algorithms for AutoML methods to select from. The potential solutions need to accept input spatial data relevant to the defined spatial problems, while allowing the chosen metrics to measure outputs in a manner that is comparable across different potential solutions. Thus, a few considerations for potential solutions include [104,160,161]:

- **Data Size**: how large or small the data are
- **Interpretability**: whether the potential solutions need to be interpreted or simply produce outputs to be used (e.g. identifying important variables versus prediction performance)
- **Resource Constraints**: time, computation, and expertise constraints (e.g. runtime of models, training to interpret results)
- **Update Frequency**: how often potential solutions need to be re-evaluated (e.g. new input data, new model/algorithm adjustments)

Similar to choosing metrics, the potential solutions generally have their own caveats and advantages [162]. For example, the performance of neural networks are reliant on the design of the neural architecture [146], while unsupervised models tend to perform better on larger datasets [161]. However, the difference between determining potential solutions and metrics is that, given the appropriate metric and adequate computing power, a comphrehensive search space can be defined and the choice of potential solutions become more flexible [163]. In more complex cases, potential solutions may also be allowed to be combined to form new potential solutions [164].

## *5.2. Implementation Challenges*

### 5.2.1. Data Quality

SDSS and AutoML rely on data for producing models to generate useful information. Data quality is an important factor in modelling as it determines whether the data are appropriate for SDSS purposes or AutoML modelling. Data contains noise [12] (e.g. errors, incompleteness), varying level of detail [165] (e.g. aggregate, scale), and risks [166] (e.g. misuse, ethics). As about 60% of the time is spent on data preparation [8], the challenge is to distribute resources to ensure that the data used is diverse (e.g. inclusive, transparent), representative (e.g. adequate coverage and detail), and reliable (e.g. minimal errors) for the intended purposes over time [23].

### 5.2.2. Model Interpretability

Predictions or actions from models are eventually explained to non-technical users in decision making (e.g. clients, society), which improves trust, transparency, and fairness [67]. Many reviewed AutoML studies measure model performance, but often do not focus on interpretability — why models perform better/worse or take certain actions [12,23]. Models in reviewed SDSS studies consider interpretability, but are often too complex for decision makers to use or communicate to stakeholders [21]. Improving interpretability leads to higher user adoption due to trust and useful knowledge from better communication. Another challenge is to balance available resources, model performance, and model interpretability to particular spatial problems.

### 5.2.3. Evidence of Usefulness

Evidence of usefulness is often not focused on in the reviewed SDSS and AutoML studies. Without measures/examinations of usefulness, it is difficult to prove the added value of SDSS or AutoML in practice. This lead to issues such as difficulty differentiating SDSS implementations [6], low SDSS user adoption [16], inconsistent AutoML performance [**?** ], and non-reusable AutoML models [64]. An important challenge is to design methods to measure practical success by examining the utility of SDSS/AutoML implementations to real world decisions — evaluating not only performance, but how SDSS with AutoML directly affect decision making.

 doi:10.20944/preprints202209.0465.v1



*5.3. Research Opportunities*

This section discusses research opportunities to address challenges in Section 5.2, where each opportunity covers one or more of the challenges (Figure 9).

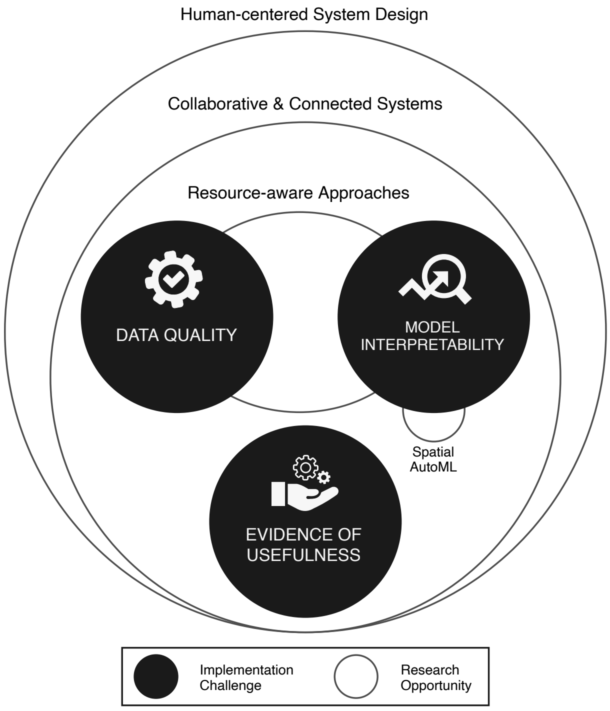


**Figure 9.** SDSS with AutoML Research Opportunities and Implementation Challenges.

5.3.1. Spatial AutoML

Many common AutoML approaches in the reviewed studies do not consider spatial data patterns when dealing with spatial problems. If spatial patterns exist in the data (e.g. clustering and dispersion in space), then the assumption that observations are independent of each other is violated and the data has spatial dependency [167]. Spatial strategies (e.g. spatial sampling [168], localized models [169]) can be integrated into AutoML approaches for potential performance or efficiency gains, while spatial parameter problems (e.g. selecting neighbours or distance bands [170]) can fit in many AutoML optimization methods. One research opportunity involves incorporating spatially explicit approaches in AutoML for SDSS to improve modelling performance/efficiency and reduce arbitrary parameter selection without strongly affecting interpretability [22].

5.3.2. Resource-aware Approaches

As SDSS and AutoML use various modelling and visualization approaches, the resources available (e.g. amount/quality of data, server processing power, data/domain experts) determine whether an implementation is feasible and practical for its intended purposes. In AutoML, the design of the search space (e.g. range of models/parameters) and optimization stopping criteria (e.g. iteration limits, reaching desired performance) are dependent on time tolerance, data available, and computing power [15]. Similarly, SDSS are dependent on hardware and software resources, but also include human resources (e.g. software developers, decision makers, consultants) that communicate and collaborate to design system features and purposes [36]. Another opportunity is to develop resource-aware implementation approaches for SDSS with AutoML by balancing the available resources and the desired results (e.g. data quality, model interpretability/performance).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1



#### 5.3.3. Collaborative and Connected Systems

Many SDSS are difficult to reuse due to being developed for specific domain purposes [16], while AutoML models are difficult to reproduce due to varying search spaces and optimization approaches [12]. Reusability/reproducibility problems hinder collaboration and communication as SDSS implementations and AutoML models do not follow standards for comparability (e.g. benchmarks, performance, features) and transferability (e.g. reuse for similar problems/different geographic area) [6,24,65,171]. Standardization of SDSS implementations and AutoML models enable SDSS with AutoML to be more easily compared/applied across various spatial problems [4], and shared across studies and organizations [36]. These standards enable different SDSS with AutoML implementations to be connected, which improves data and research transparency, availability, and interoperability (e.g. web platforms [172], programming interfaces [173], open data [38]). A third opportunity involves developing interoperable standards to connect SDSS with AutoML implementations in a network for sharing data/information/knowledge to reduce redundancy and repetition, while improving useability and collaboration between stakeholders, decision makers, and other actors.

#### 5.3.4. Human-centered System Design

Usability and interpretability are often overlooked with SDSS and AutoML research focusing on performance. When inputs and outputs are complex, barriers to translating information for useful decision making knowledge hinder non-technical users (e.g. decision makers, stakeholders, policy makers) [174]. However, a tradeoff between simplicity and accuracy exists, where improving the ease of use and transparency while minimizing complexity is desired [36,175]. SDSS with AutoML also need to balance domain-specificity (e.g. custom solutions tailored to a particular spatial problem) and adaptability (e.g. flexible solutions applicable to a variety of spatial problems) [16]. A final opportunity is to strive towards human-centered system design principles and co-design for SDSS with AutoML, where careful considerations are made regarding user and spatial problem characteristics, complexities, and interactions to enhance user adoption, experience, and practical usefulness.

## 6. Conclusion

This paper examined recent research for SDSS with AutoML, which can lower technical barriers and resource consumption to improve accessibility and adoption of SDSS for decision makers. A general framework for SDSS with AutoML was proposed by identifying and connecting general SDSS and AutoML components from select research papers, where SDSS automatically processes data into information by solving spatial problems. Challenges were discussed regarding data quality, model interpretability, and evidence of usefulness. When implementing SDSS with AutoML, there is a distribution of available resources to maintain data with adequate quality/quantity for decision making purposes, while ensuring models and systems are interpretable, comparable, and perform well in practice. Several research opportunities were also discussed to address these challenges. One opportunity involves incorporating spatially explicit models, commonly used by SDSS, in AutoML research to help optimize, standardize, and compare models used in SDSS. Other opportunities involve developing standards, approaches, and principles for resource-aware, collaborative/connected, and human-centered systems. These developments support the goal of SDSS with AutoML, which is to aid decision making involving collaboration among various actors and different resource settings. As human-related (e.g. interpretability, usability, usefulness) and technical (e.g. reproducibility, reusability, comparability) issues arise in recent SDSS and AutoML research, integrating SDSS with AutoML incorporates technical aspects of AutoML (e.g. standardized pipelines/metrics) in SDSS research, while also incorporating human-related considerations of SDSS (e.g. solution complexity, scenario evaluation) in AutoML research. SDSS with AutoML not only helps improve SDSS user

 doi:10.20944/preprints202209.0465.v1



adoption, but mutually benefits SDSS and AutoML research by fostering approaches that consider both human-related and technical issues.

**Author Contributions:** Conceptualization, methodology, investigation, data curation, software, writing—original draft preparation, R.W.; Supervision, funding acquisition, resources, project administration, writing—review & editing, S.L.

**Funding:** This work was funded by the Natural Sciences and Engineering Research Council of Canada grant number RGPIN-2017-05950.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| ABM | Agent Based Modelling |
| AI | Artificial Intelligence |
| AUC | Area Under the ROC Curve |
| AutoML | Automated Machine Learning |
| CA | Cellular Automata |
| DBSCAN | Density-Based Spatial Clustering of Applications with Noise |
| DT | Decision Trees |
| GA | Genetic Algorithms |
| GIS | Geographic Information Systems |
| GWR | Geographically Weighted Regression |
| LISA | Local Indicators of Spatial Association |
| LSTM | Long Short Term Memory |
| MAE | Mean Absolute Error |
| MCDA | Multiple Criteria Decision Analysis |
| ML | Machine Learning |
| MSE | Mean Squared Error |
| NAS | Neural Architecture Search |
| NB | Naïve Bayes |
| NN | Neural Networks |
| PSO | Particle Swarm Optimization |
| PSS | Planning Support Systems |
| RL | Reinforcement Learning |
| RMSE | Root Mean Square Error |
| SDSS | Spatial Decision Support Systems |
| SSE | Sum of Square Error |
| SVM | Support Vector Machines |
| TPOT | Tree-based Pipeline Optimization Tool |

## References

1. Niu, H.; Silva, E.A. Crowdsourced Data Mining for Urban Activity: Review of Data Sources, Applications, and Methods. *Journal of Urban Planning and Development* **2020**, *146*, 04020007. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000566.
2. Ruijer, E.; Meijer, A. Open Government Data as an Innovation Process: Lessons from a Living Lab Experiment. *Public Performance & Management Review* **2020**, *43*, 613–635. https://doi.org/10.1080/15309576.2019.1568884.
3. Riehle, D. The Innovations of Open Source. *Computer* **2019**, *52*, 59–63. https://doi.org/10.1109/MC.2019.2898163.
4. Keenan, P.B.; Jankowski, P. Spatial Decision Support Systems: Three Decades On. *Decision Support Systems* **2019**, *116*, 64–76. https://doi.org/10.1016/j.dss.2018.10.010.
5. Geertman, S. PSS: Beyond the Implementation Gap. *Transportation Research Part A: Policy and Practice* **2017**, *104*, 70–76. https://doi.org/10.1016/j.tra.2016.10.016.
6. Jiang, H.; Geertman, S.; Witte, P. Avoiding the Planning Support System Pitfalls? What Smart Governance Can Learn from the Planning Support System Implementation Gap. *Environment and Planning B: Urban Analytics and City Science* **2020**. https://doi.org/10.1177/2399808320934824.
7. Yao, Q.; Wang, M.; Chen, Y.; Dai, W.; Li, Y.F.; Tu, W.W.; Yang, Q.; Yu, Y. Taking Human out of Learning Applications: A Survey on Automated Machine Learning. *arXiv:1810.13306 [cs, stat]* **2019**, [arXiv:cs, stat/1810.13306].
8. Munson, M.A. A Study on the Importance of and Time Spent on Different Modeling Steps. *ACM SIGKDD Explorations Newsletter* **2012**, *13*, 65–71. https://doi.org/10.1145/2207243.2207253.
9. Google LLC. Cloud AutoML - Custom Machine Learning Models. https://cloud.google.com/automl, 2020.
10. Microsoft Corporation. Automated Machine Learning | Microsoft Azure. https://azure.microsoft.com/en-ca/services/machine-learning/automatedml/, 2020.
11. Amazon.com, Inc. Amazon SageMaker. https://aws.amazon.com/sagemaker/, 2020.
12. He, X.; Zhao, K.; Chu, X. AutoML: A Survey of the State-of-the-Art. *Knowledge-Based Systems* **2021**, *212*, 106622. https://doi.org/10.1016/j.knosys.2020.106622.
13. Budjač, R.; Nikmon, M.; Schreiber, P.; Zahradníková, B.; Janáčová, D. Automated Machine Learning Overview. *Vedecké Práce Materiálovotechnologickej Fakulty Slovenskej Technickej Univerzity v Bratislave so Sídlom v Trnave; Bratislava* **2019**, *27*, 107–112. https://doi.org/http://dx.doi.org.ezproxy.lib.ryerson.ca/10.2478/rput-2019-0033.
14. Weng, Z. From Conventional Machine Learning to AutoML. *Journal of Physics: Conference Series* **2019**, *1207*, 012015. https://doi.org/10.1088/1742-6596/1207/1/012015.
15. Chen, Y.W.; Song, Q.; Hu, X. Techniques for Automated Machine Learning. *ACM SIGKDD Explorations Newsletter* **2021**, *22*, 35–50. https://doi.org/10.1145/3447556.3447567.
16. Geertman, S.; Stillwell, J. Planning Support Science: Developments and Challenges:. *Environment and Planning B: Urban Analytics and City Science* **2020**. https://doi.org/10.1177/2399808320936277.
17. Flacke, J.; Shrestha, R.; Aguilar, R. Strengthening Participation Using Interactive Planning Support Systems: A Systematic Review. *ISPRS International Journal of Geo-Information* **2020**, *9*, 49. https://doi.org/10.3390/ijgi9010049.
18. Pan, H.; Geertman, S.; Deal, B. What Does Urban Informatics Add to Planning Support Technology? *Environment and Planning B: Urban Analytics and City Science* **2020**, *47*, 1317–1325. https://doi.org/10.1177/2399808320945453.
19. Lock, O.; Bain, M.; Pettit, C. Towards the Collaborative Development of Machine Learning Techniques in Planning Support Systems – a Sydney Example. *Environment and Planning B: Urban Analytics and City Science* **2020**, p. 2399808320939974. https://doi.org/10.1177/2399808320939974.
20. Niazi, M. Do Systematic Literature Reviews Outperform Informal Literature Reviews in the Software Engineering Domain? An Initial Case Study. *Arabian Journal for Science and Engineering* **2015**, *40*, 845–855. https://doi.org/10.1007/s13369-015-1586-0.
21. Pan, H.; Deal, B. Reporting on the Performance and Usability of Planning Support Systems—Towards a Common Understanding. *Applied Spatial Analysis and Policy* **2020**, *13*, 137–159. https://doi.org/10.1007/s12061-019-09296-5.
22. Du, P.; Bai, X.; Tan, K.; Xue, Z.; Samat, A.; Xia, J.; Li, E.; Su, H.; Liu, W. Advances of Four Machine Learning Methods for Spatial Data Handling: A Review. *Journal of Geovisualization and Spatial Analysis* **2020**, *4*, 13. https://doi.org/10.1007/s41651-020-00048-5.
23. Waring, J.; Lindvall, C.; Umeton, R. Automated Machine Learning: Review of the State-of-the-Art and Opportunities for Healthcare. *Artificial Intelligence in Medicine* **2020**, *104*, 101822. https://doi.org/10.1016/j.artmed.2020.101822.
24. Zöller, M.A.; Huber, M.F. Benchmark and Survey of Automated Machine Learning Frameworks. *Journal of Artificial Intelligence Research* **2021**, *70*, 409–472. https://doi.org/10.1613/jair.1.11854.
25. Taylor, R. Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonography* **1990**, *6*, 35–39. https://doi.org/10.1177/875647939000600106.
26. Oliver, M.A.; Webster, R. Kriging: A Method of Interpolation for Geographical Information Systems. *International Journal of Geographical Information Systems* **1990**, *4*, 313–332. https://doi.org/10.1080/02693799008941549.
27. Breiman, L. Random Forests. *Machine Learning* **2001**, *45*, 5–32. https://doi.org/10.1023/A:1010933404324.
28. Noble, W.S. What Is a Support Vector Machine? *Nature Biotechnology* **2006**, *24*, 1565–1567. https://doi.org/10.1038/nbt1206-1565.
29. Steinley, D. K-Means Clustering: A Half-Century Synthesis. *British Journal of Mathematical and Statistical Psychology* **2006**, *59*, 1–34. https://doi.org/10.1348/000711005X48266.
30. Sun, A.Y.; Scanlon, B.R.; Save, H.; Rateb, A. Reconstruction of GRACE Total Water Storage Through Automated Machine Learning. *Water Resources Research* **2021**, *57*, e2020WR028666. https://doi.org/10.1029/2020WR028666.

31. Babaeian, E.; Paheding, S.; Siddique, N.; Devabhaktuni, V.K.; Tuller, M. Estimation of Root Zone Soil Moisture from Ground and Remotely Sensed Soil Information with Multisensor Data Fusion and Automated Machine Learning. *Remote Sensing of Environment* **2021**, *260*, 112434. https://doi.org/10.1016/j.rse.2021.112434.
32. Singh, A.; link will open in a new window Link to external site, t.; Kumar, G.; link will open in a new window Link to external site, t.; Rai, A.K.; link will open in a new window Link to external site, t.; Beg, Z.; link will open in a new window Link to external site, t. Machine Learning to Estimate Surface Roughness from Satellite Images. *Remote Sensing* **2021**, *13*, 3794. https://doi.org/10.3390/rs13193794.
33. Hopkins, L.D.; Armstrong, M.P. Analytic and Cartographic Data Storage: A Two-Tiered Approach to Spatial Decision Support Systems. In Proceedings of the Proceedings, Seventh International Symposium on Computer-Assisted Cartography. Washington, DC: American Congress on Surveying and Mapping, 1985.
34. Longley, P.A.; Goodchild, M.F.; Maguire, D.J.; Rhind, D.W. *Geographic Information Systems and Science*; John Wiley & Sons, 2005.
35. Geertman, S.; Stillwell, J. *Planning Support Systems in Practice*; Springer Science & Business Media, 2012.
36. Schindler, M.; Dionisio, R.; Kingham, S. Challenges of Spatial Decision-Support Tools in Urban Planning: Lessons from New Zealand's Cities. *Journal of Urban Planning and Development* **2020**, *146*, 04020012. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000575.
37. Alva, P.; Janssen, P.; Stouffs, R. Geospatial Tool-Chains: Planning Support Systems for Organisational Teams. *International Journal of Architectural Computing* **2019**, *17*, 336–356. https://doi.org/10.1177/1478077119891569.
38. Zhang, G.; Zhang, W.; Guhathakurta, S.; Botchwey, N. Development of a Flow-Based Planning Support System Based on Open Data for the City of Atlanta. *Environment and Planning B: Urban Analytics and City Science* **2019**, *46*, 207–224. https://doi.org/10.1177/2399808317705881.
39. Greene, R.; Devillers, R.; Luther, J.E.; Eddy, B.G. GIS-Based Multiple-Criteria Decision Analysis. *Geography Compass* **2011**, *5*, 412–432. https://doi.org/10.1111/j.1749-8198.2011.00431.x.
40. Getis, A.; Ord, J.K. The Analysis of Spatial Association by Use of Distance Statistics. In *Perspectives on Spatial Data Analysis*; Anselin, L.; Rey, S.J., Eds.; Advances in Spatial Science, Springer: Berlin, Heidelberg, 2010; pp. 127–145. https://doi.org/10.1007/978-3-642-01976-0_10.
41. Ward, M.D.; Gleditsch, K.S. *Spatial Regression Models*; SAGE Publications, 2018.
42. Itami, R.M. Simulating Spatial Dynamics: Cellular Automata Theory. *Landscape and Urban Planning* **1994**, *30*, 27–47. https://doi.org/10.1016/0169-2046(94)90065-5.
43. Crooks, A.; Castle, C.; Batty, M. Key Challenges in Agent-Based Modelling for Geo-Spatial Simulation. *Computers, Environment and Urban Systems* **2008**, *32*, 417–430. https://doi.org/10.1016/j.compenvurbsys.2008.09.004.
44. Wahab, M.N.A.; Nefti-Meziani, S.; Atyabi, A. A Comprehensive Review of Swarm Optimization Algorithms. *PLOS ONE* **2015**, *10*, e0122827. https://doi.org/10.1371/journal.pone.0122827.
45. Shrestha, R.; Flacke, J. Leveraging Citizen Science to Advance Interactive Spatial Decision Support Technology: A Swot Analysis. In Proceedings of the The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences; Gottingen; Copernicus GmbH: Gottingen, Germany, Gottingen, 2019; Vol. XLII-5/W3, pp. 99–105. https://doi.org/http://dx.doi.org.ezproxy.lib.ryerson.ca/10.5194/isprs-archives-XLII-5-W3-99-2019.
46. Mutuku, B.; Boerboom, L.; Madureira, A.M. The Role of Planning Support Systems in National Policy Transfer and Policy Translation in Secondary Cities. *International Planning Studies* **2019**, *24*, 293–307. https://doi.org/10.1080/13563475.2019.1657809.
47. Erskine, M.A.; Gregg, D.G.; Karimi, J.; Scott, J.E. Individual Decision-Performance Using Spatial Decision Support Systems: A Geospatial Reasoning Ability and Perceived Task-Technology Fit Perspective. *Information Systems Frontiers* **2019**, *21*, 1369–1384. https://doi.org/10.1007/s10796-018-9840-0.
48. Punt, E.P.; Geertman, S.C.M.; Afrooz, A.E.; Witte, P.A.; Pettit, C.J. Life Is a Scene and We Are the Actors: Assessing the Usefulness of Planning Support Theatres for Smart City Planning. *Computers, Environment and Urban Systems* **2020**, *82*, 101485. https://doi.org/10.1016/j.compenvurbsys.2020.101485.
49. Maceachren, A.M.; Brewer, I. Developing a Conceptual Framework for Visually-Enabled Geocollaboration. *International Journal of Geographical Information Science* **2004**, *18*, 1–34. https://doi.org/10.1080/13658810310001596094.
50. Page, J.; Mörtberg, U.; Destouni, G.; Ferreira, C.; Nässtöm, H.; Kalantari, Z. Open-Source Planning Support System for Sustainable Regional Planning: A Case Study of Stockholm County, Sweden:. *Environment and Planning B: Urban Analytics and City Science* **2020**. https://doi.org/10.1177/2399808320919769.
51. Hooper, P.; Boulange, C.; Arciniegas, G.; Foster, S.; Bolleter, J.; Pettit, C. Exploring the Potential for Planning Support Systems to Bridge the Research-Translation Gap between Public Health and Urban Planning. *International Journal of Health Geographics* **2021**, *20*, 36. https://doi.org/10.1186/s12942-021-00291-z.
52. Daniel, C.; Pettit, C. Charting the Past and Possible Futures of Planning Support Systems: Results of a Citation Network Analysis. *Environment and Planning B: Urban Analytics and City Science* **2022**, *49*, 1875–1892. https://doi.org/10.1177/23998083211072866.
53. Golnaraghi, F.; Kuo, B.C. *Automatic Control Systems*, 9th edition ed.; Wiley: Hoboken, NJ, 2009.
54. Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development* **1959**, *3*, 210–229. https://doi.org/10.1147/rd.33.0210.
55. Seber, G.A.F.; Lee, A.J. *Linear Regression Analysis*; John Wiley & Sons, 2012.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1




56. Kleinbaum, D.G.; Klein, M. Introduction to Logistic Regression. In *Logistic Regression: A Self-Learning Text*; Kleinbaum, D.G.; Klein, M., Eds.; Statistics for Biology and Health, Springer: New York, NY, 2010; pp. 1–39. https://doi.org/10.1007/978-1-4419-1742-3_1.
57. Rish, I. An Empirical Study of the Naive Bayes Classifier. Technical report, 2001.
58. Quinlan, J.R. Decision Trees and Decision-Making. *IEEE Transactions on Systems, Man, and Cybernetics* **1990**, *20*, 339–346. https://doi.org/10.1109/21.52545.
59. Hinton, G.E. How Neural Networks Learn from Experience. *Scientific American* **1992**, *267*, 144–151.
60. Jaramillo, J.H.; Bhadury, J.; Batta, R. On the Use of Genetic Algorithms to Solve Location Problems. *Computers & Operations Research* **2002**, *29*, 761–779. https://doi.org/10.1016/S0305-0548(01)00021-1.
61. Santu, S.K.K.; Hassan, M.M.; Smith, M.J.; Xu, L.; Zhai, C.; Veeramachaneni, K. A Level-wise Taxonomic Perspective on Automated Machine Learning to Date and Beyond: Challenges and Opportunities. *arXiv:2010.10777 [cs]* **2020**, [arXiv:cs/2010.10777].
62. Feurer, M.; Klein, A.; Eggensperger, K.; Springenberg, J.T.; Blum, M.; Hutter, F. Auto-Sklearn: Efficient and Robust Automated Machine Learning. In *Automated Machine Learning*; Hutter, F.; Kotthoff, L.; Vanschoren, J., Eds.; Springer International Publishing: Cham, 2019; pp. 113–134. https://doi.org/10.1007/978-3-030-05318-5_6.
63. Escalante, H.J.; Tu, W.W.; Guyon, I.; Silver, D.L.; Viegas, E.; Chen, Y.; Dai, W.; Yang, Q. AutoML @ NeurIPS 2018 Challenge: Design and Results. In Proceedings of the The NeurIPS '18 Competition; Escalera, S.; Herbrich, R., Eds.; Springer International Publishing: Cham, 2020; The Springer Series on Challenges in Machine Learning, pp. 209–229. https://doi.org/10.1007/978-3-030-29135-8_8.
64. Madrid, J.G.; Jair Escalante, H.; Morales, E.F.; Tu, W.W.; yu, y.; Sun-Hosoya, L.; Guyon, I.; Sebag, M. Towards AutoML in the Presence of Drift: First Results. In Proceedings of the Workshop AutoML 2018 @ ICML/IJCAI-ECAI; Pavel Brazdil and Christophe Giraud-Carrier and Isabelle Guyon: Stockholm, Sweden, 2018.
65. Halvari, T.; Nurminen, J.K.; Mikkonen, T. Testing the Robustness of AutoML Systems. *Electronic Proceedings in Theoretical Computer Science* **2020**, *319*, 103–116, [2005.02649]. https://doi.org/10.4204/EPTCS.319.8.
66. Molnar, C. *Interpretable Machine Learning*; Lulu.com, 2020.
67. Pfisterer, F.; Thomas, J.; Bischl, B. Towards Human Centered AutoML. *arXiv:1911.02391 [cs]* **2019**, [arXiv:cs/1911.02391].
68. Karmaker, S.K.; Hassan, M.M.; Smith, M.J.; Xu, L.; Zhai, C.; Veeramachaneni, K. AutoML to Date and Beyond: Challenges and Opportunities. *ACM Computing Surveys* **2021**, *54*, 175:1–175:36. https://doi.org/10.1145/3470918.
69. Bahri, M.; Salutari, F.; Putina, A.; Sozio, M. AutoML: State of the Art with a Focus on Anomaly Detection, Challenges, and Research Directions. *International Journal of Data Science and Analytics* **2022**, *14*, 113–126. https://doi.org/10.1007/s41060-022-00309-0.
70. Alsharef, A.; Aggarwal, K.; Sonia.; Kumar, M.; Mishra, A. Review of ML and AutoML Solutions to Forecast Time-Series Data. *Archives of Computational Methods in Engineering* **2022**. https://doi.org/10.1007/s11831-022-09765-0.
71. Li, W. GeoAI: Where Machine Learning and Big Data Converge in GIScience. *Journal of Spatial Information Science* **2020**, pp. 71–77. https://doi.org/10.5311/JOSIS.2020.20.658.
72. Janowicz, K.; Gao, S.; McKenzie, G.; Hu, Y.; Bhaduri, B. GeoAI: Spatially Explicit Artificial Intelligence Techniques for Geographic Knowledge Discovery and Beyond. *International Journal of Geographical Information Science* **2020**, *34*, 625–636. https://doi.org/10.1080/13658816.2019.1684500.
73. Fang, Z.; Jin, Y.; Yang, T. Incorporating Planning Intelligence into Deep Learning: A Planning Support Tool for Street Network Design. *Journal of Urban Technology* **2022**, *29*, 99–114. https://doi.org/10.1080/10630732.2021.2001713.
74. Myers, D.E. Spatial Interpolation: An Overview. *Geoderma* **1994**, *62*, 17–28. https://doi.org/10.1016/0016-7061(94)90025-6.
75. Jiang, Z. A Survey on Spatial Prediction Methods. *IEEE Transactions on Knowledge and Data Engineering* **2019**, *31*, 1645–1664. https://doi.org/10.1109/TKDE.2018.2866809.
76. Unwin, D. Integration through Overlay Analysis. In *Spatial Analytical Perspectives on GIS*; Routledge, 1996.
77. Das, S.; Li, J.J.; Allston, A.; Kharfen, M. Planning Area-Specific Prevention and Intervention Programs for HIV Using Spatial Regression Analysis. *Public Health* **2019**, *169*, 41–49. https://doi.org/10.1016/j.puhe.2019.01.009.
78. Costache, R.; Popa, M.C.; Tien Bui, D.; Diaconu, D.C.; Ciubotaru, N.; Minea, G.; Pham, Q.B. Spatial Predicting of Flood Potential Areas Using Novel Hybridizations of Fuzzy Decision-Making, Bivariate Statistics, and Machine Learning. *Journal of Hydrology* **2020**, *585*, 124808. https://doi.org/10.1016/j.jhydrol.2020.124808.
79. Warth, G.; Braun, A.; Assmann, O.; Fleckenstein, K.; Hochschild, V. Prediction of Socio-Economic Indicators for Urban Planning Using VHR Satellite Imagery and Spatial Analysis. *Remote Sensing* **2020**, *12*, 1730. https://doi.org/10.3390/rs12111730.
80. Brunsdon, C.; Fotheringham, S.; Charlton, M. Geographically Weighted Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)* **1998**, *47*, 431–443. https://doi.org/10.1111/1467-9884.00145.
81. Naser, M.Z.; Alavi, A.H. Error Metrics and Performance Fitness Indicators for Artificial Intelligence and Machine Learning in Engineering and Sciences. *Architecture, Structures and Construction* **2021**. https://doi.org/10.1007/s44150-021-00015-8.
82. Bruno, G.; Giannikos, I. Location and GIS. In *Location Science*; Laporte, G.; Nickel, S.; Saldanha da Gama, F., Eds.; Springer International Publishing: Cham, 2015; pp. 509–536. https://doi.org/10.1007/978-3-319-13111-5_19.
83. Keenan, P.B. Spatial Decision Support Systems for Vehicle Routing. *Decision Support Systems* **1998**, *22*, 65–71. https://doi.org/10.1016/S0167-9236(97)00054-7.
84. Farahani, R.Z.; Asgari, N.; Heidari, N.; Hosseininia, M.; Goh, M. Covering Problems in Facility Location: A Review. *Computers & Industrial Engineering* **2012**, *62*, 368–407. https://doi.org/10.1016/j.cie.2011.08.020.
85. Keenan, P. Modelling Vehicle Routing in GIS. *Operational Research* **2008**, *8*, 201. https://doi.org/10.1007/s12351-008-0021-7.

86. Laporte, G.; Martello, S. The Selective Travelling Salesman Problem. *Discrete Applied Mathematics* **1990**, *26*, 193–207. https://doi.org/10.1016/0166-218X(90)90100-Q.
87. Kaveh, M.; Kaveh, M.; Mesgari, M.S.; Paland, R.S. Multiple Criteria Decision-Making for Hospital Location-Allocation Based on Improved Genetic Algorithm. *Applied Geomatics* **2020**. https://doi.org/10.1007/s12518-020-00297-5.
88. Diemuodeke, E.O.; Addo, A.; Oko, C.O.C.; Mulugetta, Y.; Ojapah, M.M. Optimal Mapping of Hybrid Renewable Energy Systems for Locations Using Multi-Criteria Decision-Making Algorithm. *Renewable Energy* **2019**, *134*, 461–477. https://doi.org/10.1016/j.renene.2018.11.055.
89. Musolino, G.; Rindone, C.; Polimeni, A.; Vitetta, A. Planning Urban Distribution Center Location with Variable Restocking Demand Scenarios: General Methodology and Testing in a Medium-Size Town. *Transport Policy* **2019**, *80*, 157–166. https://doi.org/10.1016/j.tranpol.2018.04.006.
90. Wang, Z.; Li, M.; Tang, L.; Huang, J.S. Research and Application of Intersection Traffic Signal Control Algorithm Based on Vehicle Location. *International Journal of Communication Networks and Distributed Systems* **2020**, *24*, 249–261. https://doi.org/10.1504/IJCNDS.2020.106355.
91. de Montis, A.; Toro, P.D.; Droste-Franke, B.; Stagl, I.O.a.S. Assessing the Quality of Different MCDA Methods. In *Alternatives for Environmental Valuation*; Routledge, 2004.
92. Riquelme, N.; Von Lücken, C.; Baran, B. Performance Metrics in Multi-Objective Optimization. In Proceedings of the 2015 Latin American Computing Conference (CLEI), 2015, pp. 1–11. https://doi.org/10.1109/CLEI.2015.7360024.
93. Li, X.; Yeh, A.G.O. Integration of Genetic Algorithms and GIS for Optimal Location Search. *International Journal of Geographical Information Science* **2005**, *19*, 581–601. https://doi.org/10.1080/13658810500032388.
94. Anselin, L. Local Indicators of Spatial Association—LISA. *Geographical Analysis* **1995**, *27*, 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.
95. Khan, K.; Rehman, S.U.; Aziz, K.; Fong, S.; Sarasvady, S. DBSCAN: Past, Present and Future. In Proceedings of the The Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014), 2014, pp. 232–238. https://doi.org/10.1109/ICADIWT.2014.6814687.
96. Ansari, M.Y.; Ahmad, A.; Khan, S.S.; Bhushan, G.; Mainuddin. Spatiotemporal Clustering: A Review. *Artificial Intelligence Review* **2020**, *53*, 2381–2423. https://doi.org/10.1007/s10462-019-09736-1.
97. Kulldorff, M. A Spatial Scan Statistic. *Communications in Statistics - Theory and Methods* **1997**, *26*, 1481–1496. https://doi.org/10.1080/03610929708831995.
98. Aldstadt, J. Spatial Clustering. In *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*; Fischer, M.M.; Getis, A., Eds.; Springer: Berlin, Heidelberg, 2010; pp. 279–300. https://doi.org/10.1007/978-3-642-03647-7_15.
99. Irandegani, Z.; Mohammadi, R.; Taleai, M. Investigating Temporal and Spatial Effects of Urban Planning Variables on Crime Rate: A Gwr and Ols Based Approach. In Proceedings of the The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences; Gottingen; Copernicus GmbH: Gottingen, Germany, Gottingen, 2019; Vol. XLII-4/W18, pp. 559–564. https://doi.org/http://dx.doi.org.ezproxy.lib.ryerson.ca/10.5194/isprs-archives-XLII-4-W18-559-2019.
100. Ohana-Levi, N.; Ben-Gal, A.; Peeters, A.; Termin, D.; Linker, R.; Baram, S.; Raveh, E.; Paz-Kagan, T. A Comparison between Spatial Clustering Models for Determining N-fertilization Management Zones in Orchards. *Precision Agriculture* **2020**. https://doi.org/10.1007/s11119-020-09731-5.
101. Fitzmaurice, A.G.; Linley, L.; Zhang, C.; Watson, M.; France, A.M.; Oster, A.M. Novel Method for Rapid Detection of Spatiotemporal HIV Clusters Potentially Warranting Intervention. *Emerging Infectious Diseases* **2019**, *25*, 988–991. https://doi.org/10.3201/eid2505.180776.
102. Li, M.; Croitoru, A.; Yue, S. GeoDenStream: An Improved DenStream Clustering Method for Managing Entity Data within Geographical Data Streams. *Computers & Geosciences* **2020**, *144*, 104563. https://doi.org/10.1016/j.cageo.2020.104563.
103. Peterson, B.A.; Brownlee, M.T.J.; Hallo, J.C.; Beeco, J.A.; White, D.L.; Sharp, R.L.; Cribbs, T.W. Spatiotemporal Variables to Understand Visitor Travel Patterns: A Management-Centric Approach. *Journal of Outdoor Recreation and Tourism* **2020**, *31*, 100316. https://doi.org/10.1016/j.jort.2020.100316.
104. Grubesic, T.H.; Wei, R.; Murray, A.T. Spatial Clustering Overview and Comparison: Accuracy, Sensitivity, and Computational Expense. *Annals of the Association of American Geographers* **2014**, *104*, 1134–1156. https://doi.org/10.1080/00045608.2014.958389.
105. Grabusts, P. The Choice of Metrics for Clustering Algorithms. *ENVIRONMENT. TECHNOLOGIES. RESOURCES. Proceedings of the International Scientific and Practical Conference* **2011**, *2*, 70–76. https://doi.org/10.17770/etr2011vol2.973.
106. Surendran, M.S. Review of Spatial Clustering Methods. *International Journal of Information Technology Infrastructure* **2012**, *2*, 15–24.
107. Fritz, C.E.; Schuurman, N.; Robertson, C.; Lear, S. A Scoping Review of Spatial Cluster Analysis Techniques for Point-Event Data. *Geospatial Health* **2013**, *7*, 183–198. https://doi.org/10.4081/gh.2013.79.
108. O'Sullivan, D.; Perry, G.L.W. *Spatial Simulation: Exploring Pattern and Process*; John Wiley & Sons, 2013.
109. Sinclair, T.R.; Soltani, A.; Marrou, H.; Ghanem, M.; Vadez, V. Geospatial Assessment for Crop Physiological and Management Improvements with Examples Using the Simple Simulation Model. *Crop Science* **2020**, *60*, 700–708. https://doi.org/10.1002/csc2.20106.
110. Chen, L.; Zhang, P.; Lv, G.P.; Shen, Z.Y. Spatial–Temporal Distribution and Limiting Factor Variation of Algal Growth: Three-Dimensional Simulation to Enhance Drinking Water Reservoir Management. *International Journal of Environmental Science and Technology* **2019**, *16*, 7417–7432. https://doi.org/10.1007/s13762-018-2113-0.

111. Wang, Y.; Xu, T.; Niu, X.; Tan, C.; Chen, E.; Xiong, H. STMARL: A Spatio-Temporal Multi-Agent Reinforcement Learning Approach for Cooperative Traffic Light Control. *IEEE Transactions on Mobile Computing* **2022**, *21*, 2228–2242. https://doi.org/10.1109/TMC.2020.3033782.
112. Hesam, S.; Valizadeh Kamran, K. INTELLIGENT MANAGEMENT OCCURRENCE AND SPREAD OF FRONT FIRE IN GIS BY USING CELLULAR AUTOMATA. CASE STUDY: GOLESTAN FOREST. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2019**, *XLII-4/W18*, 475–481. https://doi.org/10.5194/isprs-archives-XLII-4-W18-475-2019.
113. Yu, D.; Yanxu, L.; Bojie, F. Urban Growth Simulation Guided by Ecological Constraints in Beijing City: Methods and Implications for Spatial Planning. *Journal of Environmental Management* **2019**, *243*, 402–410. https://doi.org/10.1016/j.jenvman.2019.04.087.
114. Parker, D.C.; Meretsky, V. Measuring Pattern Outcomes in an Agent-Based Model of Edge-Effect Externalities Using Spatial Metrics. *Agriculture, Ecosystems & Environment* **2004**, *101*, 233–250. https://doi.org/10.1016/j.agee.2003.09.007.
115. Wallentin, G. Spatial Simulation: A Spatial Perspective on Individual-Based Ecology—a Review. *Ecological Modelling* **2017**, *350*, 30–41. https://doi.org/10.1016/j.ecolmodel.2017.01.017.
116. Anselin, L. Under the Hood Issues in the Specification and Interpretation of Spatial Regression Models. *Agricultural Economics* **2002**, *27*, 247–267. https://doi.org/10.1111/j.1574-0862.2002.tb00120.x.
117. Bailey, T.C. GIS and Simple Systems for Visual, Interactive, Spatial Analysis. *The Cartographic Journal* **1990**, *27*, 79–84. https://doi.org/10.1179/000870490787858676.
118. Tyner, J.A. *Principles of Map Design*; Guilford Publications, 2014.
119. Rivest, S. Toward Better Support for Spatial Decision Making: Defining the Characteristics of Spatial on-Line Analytical Processing (Solap). *GEOMATICA* **2001**, *55*, 539–555. https://doi.org/10.5623/geomat-2001-0061.
120. Kraak, J.M.; Brown, A. *Web Cartography*; CRC Press, 2003.
121. Chi, G.; Zhu, J. Spatial Regression Models for Demographic Analysis. *Population Research and Policy Review* **2008**, *27*, 17–42. https://doi.org/10.1007/s11113-007-9051-8.
122. Wu, Z.; Chen, Y.; Han, Y.; Ke, T.; Liu, Y. Identifying the Influencing Factors Controlling the Spatial Variation of Heavy Metals in Suburban Soil Using Spatial Regression Models. *Science of The Total Environment* **2020**, *717*, 137212. https://doi.org/10.1016/j.scitotenv.2020.137212.
123. Feng, Q.; Flanagan, D.C.; Engel, B.A.; Yang, L.; Chen, L. GeoAPEXOL, a Web GIS Interface for the Agricultural Policy Environmental eXtender (APEX) Model Enabling Both Field and Small Watershed Simulation. *Environmental Modelling & Software* **2020**, *123*, 104569. https://doi.org/10.1016/j.envsoft.2019.104569.
124. Ono, J.P.; Castelo, S.; Lopez, R.; Bertini, E.; Freire, J.; Silva, C. PipelineProfiler: A Visual Analytics Tool for the Exploration of AutoML Pipelines. *IEEE Transactions on Visualization and Computer Graphics* **2020**, pp. 1–1. https://doi.org/10.1109/TVCG.2020.3030361.
125. Wei, P.; Lu, Z.; Song, J. Variable Importance Analysis: A Comprehensive Review. *Reliability Engineering & System Safety* **2015**, *142*, 399–432. https://doi.org/10.1016/j.ress.2015.05.018.
126. Carvalho, D.V.; Pereira, E.M.; Cardoso, J.S. Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics* **2019**, *8*, 832. https://doi.org/10.3390/electronics8080832.
127. North, C. Toward Measuring Visualization Insight. *IEEE Computer Graphics and Applications* **2006**, *26*, 6–9. https://doi.org/10.1109/MCG.2006.70.
128. Andrienko, N.; Andrienko, G.; Gatalsky, P. Exploratory Spatio-Temporal Visualization: An Analytical Review. *Journal of Visual Languages & Computing* **2003**, *14*, 503–541. https://doi.org/10.1016/S1045-926X(03)00046-6.
129. Hallisey, E.J. Cartographic Visualization: An Assessment and Epistemological Review *. *The Professional Geographer* **2005**, *57*, 350–364. https://doi.org/10.1111/j.0033-0124.2005.00483.x.
130. Kasimati, A.; Espejo-García, B.; Darra, N.; Fountas, S. Predicting Grape Sugar Content under Quality Attributes Using Normalized Difference Vegetation Index Data and Automated Machine Learning. *Sensors (Basel, Switzerland)* **2022**, *22*, 3249. https://doi.org/10.3390/s22093249.
131. Kai-Yun, L.; link will open in a new window Link to external site, t.; Sampaio de Lima, R.; link will open in a new window Link to external site, t.; Burnside, N.G.; link will open in a new window Link to external site, t.; Vahtmäe, E.; Kutser, T.; link will open in a new window Link to external site, t.; Sepp, K.; et al. Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation. *Remote Sensing* **2022**, *14*, 1114. https://doi.org/10.3390/rs14051114.
132. Dilmurat, K.; Sagan, V.; Moose, S. Ai-Driven Maize Yield Forecasting Using Unmanned Aerial Vehicle-Based Hyperspectral and Lidar Data Fusion. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2022**, *V-3-2022*, 193–199. https://doi.org/10.5194/isprs-annals-V-3-2022-193-2022.
133. Kai-Yun, L.; link will open in a new window Link to external site, t.; Burnside, N.G.; link will open in a new window Link to external site, t.; Sampaio de Lima, R.; link will open in a new window Link to external site, t.; Peciña, M.V.; Sepp, K.; Cabral Pinheiro, V.H.; de Lima, B.R.C.A.; et al. An Automated Machine Learning Framework in Unmanned Aircraft Systems: New Insights into Agricultural Management Practices Recognition Approaches. *Remote Sensing* **2021**, *13*, 3190. https://doi.org/10.3390/rs13163190.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2022 doi:10.20944/preprints202209.0465.v1




134. Gerassis, S.; link will open in a new window Link to external site, t.; Giráldez, E.; Pazo-Rodríguez, M.; Saavedra, Á.; link will open in a new window Link to external site, t.; Taboada, J. AI Approaches to Environmental Impact Assessments (EIAs) in the Mining and Metals Sector Using AutoML and Bayesian Modeling. *Applied Sciences* **2021**, *11*, 7914. https://doi.org/10.3390/app11177914.
135. Guo, Y.; Quan, L.; Song, L.; Liang, H. Construction of Rapid Early Warning and Comprehensive Analysis Models for Urban Waterlogging Based on AutoML and Comparison of the Other Three Machine Learning Algorithms. *Journal of Hydrology* **2022**, *605*, 127367. https://doi.org/10.1016/j.jhydrol.2021.127367.
136. Bai, Z.; Liu, Q.; Liu, Y. Groundwater Potential Mapping in Hubei Region of China Using Machine Learning, Ensemble Learning, Deep Learning and AutoML Methods. *Natural Resources Research* **2022**, *31*, 2549–2569. https://doi.org/10.1007/s11053-022-10100-4.
137. Zhang, X.; Jin, Q.; Yu, T.; Xiang, S.; Kuang, Q.; Prinet, V.; Pan, C. Multi-Modal Spatio-Temporal Meteorological Forecasting with Deep Neural Network. *ISPRS Journal of Photogrammetry and Remote Sensing* **2022**, *188*, 380–393. https://doi.org/10.1016/j.isprsjprs.2022.03.007.
138. O'Donncha, F.; Hu, Y.; Palmes, P.; Burke, M.; Filgueira, R.; Grant, J. A Spatio-Temporal LSTM Model to Forecast across Multiple Temporal and Spatial Scales. *Ecological Informatics* **2022**, *69*, 101687. https://doi.org/10.1016/j.ecoinf.2022.101687.
139. Nikitin, N.O.; Revin, I.; Hvatov, A.; Vychuzhanin, P.; Kalyuzhnaya, A.V. Hybrid and Automated Machine Learning Approaches for Oil Fields Development: The Case Study of Volve Field, North Sea. *Computers & Geosciences* **2022**, *161*, 105061. https://doi.org/10.1016/j.cageo.2022.105061.
140. Bruzón, A.G.; Arrogante-Funes, P.; Arrogante-Funes, F.; Martín-González, F.; Novillo, C.J.; Fernández, R.R.; Vázquez-Jiménez, R.; Alarcón-Paredes, A.; Alonso-Silverio, G.A.; Cantu-Ramirez, C.A.; et al. Landslide Susceptibility Assessment Using an AutoML Framework. *International Journal of Environmental Research and Public Health* **2021**, *18*, 10971. https://doi.org/10.3390/ijerph182010971.
141. Arrogante-Funes, P.; Bruzón, A.G.; Arrogante-Funes, F.; Ramos-Bernal, R.N.; Vázquez-Jiménez, R. Integration of Vulnerability and Hazard Factors for Landslide Risk Assessment. *International Journal of Environmental Research and Public Health* **2021**, *18*, 11987. https://doi.org/10.3390/ijerph182211987.
142. Siriborvornratanakul, T. Human Behavior in Image-Based Road Health Inspection Systems despite the Emerging AutoML. *Journal of Big Data* **2022**, *9*, 96. https://doi.org/10.1186/s40537-022-00646-8.
143. D'Orazio, V.; Lin, Y. Forecasting Conflict in Africa with Automated Machine Learning Systems. *International Interactions* **2022**, *0*, 1–25. https://doi.org/10.1080/03050629.2022.2017290.
144. Sagi, O.; Rokach, L. Ensemble Learning: A Survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **2018**, *8*, e1249.
145. Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in neural information processing systems* **2012**, *25*.
146. Elsken, T.; Metzen, J.H.; Hutter, F. Neural Architecture Search: A Survey. *The Journal of Machine Learning Research* **2019**, *20*, 1997–2017.
147. Al-Sahaf, H.; Bi, Y.; Chen, Q.; Lensen, A.; Mei, Y.; Sun, Y.; Tran, B.; Xue, B.; Zhang, M. A Survey on Evolutionary Machine Learning. *Journal of the Royal Society of New Zealand* **2019**, *49*, 205–228.
148. Feurer, M.; Klein, A.; Eggensperger, K.; Springenberg, J.; Blum, M.; Hutter, F. Efficient and Robust Automated Machine Learning. *Advances in neural information processing systems* **2015**, *28*.
149. Olson, R.S.; Bartley, N.; Urbanowicz, R.J.; Moore, J.H. Evaluation of a Tree-Based Pipeline Optimization Tool for Automating Data Science. In Proceedings of the Proceedings of the Genetic and Evolutionary Computation Conference 2016, 2016, pp. 485–492.
150. LeDell, E.; Poirier, S. H2o Automl: Scalable Automatic Machine Learning. In Proceedings of the Proceedings of the AutoML Workshop at ICML, 2020, Vol. 2020.
151. Fakoor, R.; Mueller, J.W.; Erickson, N.; Chaudhari, P.; Smola, A.J. Fast, Accurate, and Simple Models for Tabular Data via Augmented Distillation. *Advances in Neural Information Processing Systems* **2020**, *33*, 8671–8681.
152. Drori, I.; Krishnamurthy, Y.; Rampin, R.; Lourenco, R.d.P.; Ono, J.P.; Cho, K.; Silva, C.; Freire, J. AlphaD3M: Machine Learning Pipeline Synthesis. *arXiv preprint arXiv:2111.02508* **2021**, [2111.02508].
153. Sprague, R.H. A Framework for the Development of Decision Support Systems. *MIS Quarterly* **1980**, *4*, 1–26. https://doi.org/10.2307/248957.
154. Keen, P.G.; Hackathorn, R.D. Decision Support Systems and Personal Computing **1979**.
155. Simon, H.A. *The New Science of Management Decision*; The New Science of Management Decision, Harper & Brothers: New York, NY, US, 1960; pp. xii, 50. https://doi.org/10.1037/13978-000.
156. Omran, M.G.; Engelbrecht, A.P.; Salman, A. An Overview of Clustering Methods. *Intelligent Data Analysis* **2007**, *11*, 583–605.
157. Harkanth, S.; Phulpagar, B.D. A Survey on Clustering Methods and Algorithms. *International Journal of Computer Science and Information Technologies* **2013**, *4*, 687–691.
158. Amigó, E.; Gonzalo, J.; Artiles, J.; Verdejo, F. A Comparison of Extrinsic Clustering Evaluation Metrics Based on Formal Constraints. *Information Retrieval* **2009**, *12*, 461–486. https://doi.org/10.1007/s10791-008-9066-8.
159. Maulik, U.; Bandyopadhyay, S. Performance Evaluation of Some Clustering Algorithms and Validity Indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2002**, *24*, 1650–1654. https://doi.org/10.1109/TPAMI.2002.1114856.

160. Vazquezl, M.Y.L.; Peñafiel, L.A.B.; Muñoz, S.X.S.; Martinez, M.A.Q. A Framework for Selecting Machine Learning Models Using TOPSIS. In Proceedings of the Advances in Artificial Intelligence, Software and Systems Engineering; Ahram, T., Ed.; Springer International Publishing: Cham, 2021; Advances in Intelligent Systems and Computing, pp. 119–126. https://doi.org/10.1007/978-3-030-51328-3_18.
161. Mahesh, B. Machine Learning Algorithms-a Review. *International Journal of Science and Research (IJSR).[Internet]* **2020**, *9*, 381–386.
162. Mayfield, H.J.; Smith, C.; Gallagher, M.; Hockings, M. Considerations for Selecting a Machine Learning Technique for Predicting Deforestation. *Environmental Modelling & Software* **2020**, *131*, 104741. https://doi.org/10.1016/j.envsoft.2020.104741.
163. Sparks, E.R.; Talwalkar, A.; Haas, D.; Franklin, M.J.; Jordan, M.I.; Kraska, T. Automating Model Search for Large Scale Machine Learning. In Proceedings of the Proceedings of the Sixth ACM Symposium on Cloud Computing; Association for Computing Machinery: New York, NY, USA, 2015; SoCC '15, pp. 368–380. https://doi.org/10.1145/2806777.2806945.
164. Real, E.; Liang, C.; So, D.; Le, Q. AutoML-Zero: Evolving Machine Learning Algorithms From Scratch. In Proceedings of the Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020, pp. 8007–8019.
165. Biljecki, F.; Heuvelink, G.B.M.; Ledoux, H.; Stoter, J. The Effect of Acquisition Error and Level of Detail on the Accuracy of Spatial Analyses. *Cartography and Geographic Information Science* **2018**, *45*, 156–176. https://doi.org/10.1080/15230406.2017.1279986.
166. Devillers, R.; Bédard, Y.; Jeansoulin, R.; Moulin, B. Towards Spatial Data Quality Information Analysis Tools for Experts Assessing the Fitness for Use of Spatial Data. *International Journal of Geographical Information Science* **2007**, *21*, 261–282. https://doi.org/10.1080/13658810600911879.
167. Getis, A. Spatial Autocorrelation. In *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*; Fischer, M.M.; Getis, A., Eds.; Springer: Berlin, Heidelberg, 2010; pp. 255–278. https://doi.org/10.1007/978-3-642-03647-7_14.
168. Wang, J.F.; Stein, A.; Gao, B.B.; Ge, Y. A Review of Spatial Sampling. *Spatial Statistics* **2012**, *2*, 1–14. https://doi.org/10.1016/j.spasta.2012.08.001.
169. Gilardi, N.; Bengio, S. Local Machine Learning Models for Spatial Data Analysis. *Journal of Geographic Information and Decision Analysis* **2000**, *4*, 11–28. https://doi.org/10.1.1.384.908.
170. Getis, A.; Aldstadt, J. Constructing the Spatial Weights Matrix Using a Local Statistic. *Geographical Analysis* **2004**, *36*, 90–104. https://doi.org/10.1111/j.1538-4632.2004.tb01127.x.
171. Hanussek, M.; Blohm, M.; Kintz, M. Can AutoML Outperform Humans? An Evaluation on Popular OpenML Datasets Using AutoML Benchmark. In Proceedings of the 2020 2nd International Conference on Artificial Intelligence, Robotics and Control; Association for Computing Machinery: New York, NY, USA, 2020; AIRC'20, pp. 29–32. https://doi.org/10.1145/3448326.3448353.
172. Rouse, L.J.; Bergeron, S.J.; Harris, T.M. Participating in the Geospatial Web: Collaborative Mapping, Social Networks and Participatory GIS. In *The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 Are Shaping the Network Society*; Scharl, A.; Tochtermann, K., Eds.; Advanced Information and Knowledge Processing, Springer: London, 2007; pp. 153–158. https://doi.org/10.1007/978-1-84628-827-2_14.
173. Zambelli, P.; Gebbert, S.; Ciolli, M. Pygrass: An Object Oriented Python Application Programming Interface (API) for Geographic Resources Analysis Support System (GRASS) Geographic Information System (GIS). *ISPRS International Journal of Geo-Information* **2013**, *2*, 201–219. https://doi.org/10.3390/ijgi2010201.
174. Vonk, G.; Geertman, S. Improving the Adoption and Use of Planning Support Systems in Practice. *Applied Spatial Analysis and Policy* **2008**, *1*, 153–173. https://doi.org/10.1007/s12061-008-9011-7.
175. Hong, S.R.; Castelo, S.; D'Orazio, V.; Benthune, C.; Santos, A.; Langevin, S.; Jonker, D.; Bertini, E.; Freire, J. Towards Evaluating Exploratory Model Building Process with AutoML Systems. *arXiv:2009.00449 [cs]* **2020**, [arXiv:cs/2009.00449].