Accepted Manuscript

Vietnam Journal of Computer Science

This is an unedited version of the accepted manuscript scheduled for publication. It has been uploaded in advance for the benefit of our customers. The manuscript will be copyedited, typeset and proofread before it is released in the final form. As a result, the published copy may differ from the unedited version. Readers should obtain the final version from the above link when it is published. The authors are responsible for the content of this Accepted Article.

World Scientific
www.worldscientific.com

Machine Learning Methods in BIM-based Applications – a Review

Grażyna Ślusarczyk[a]

Institute of Applied Computer Science, Jagiellonian University, Lojasiewicza 11, 30-059 Kraków, Poland
grazyna.slusarczyk@uj.edu.pl

Barbara Strug

Institute of Applied Computer Science, Jagiellonian University, Lojasiewicza 11, 30-059 Kraków, Poland
barbara.strug@uj.edu.pl

This paper presents a survey of machine learning methods used in applications dedicated to building and construction industry. A BIM model being a database system for civil engineering data is presented. A representative selection of methods and applications is described. The aim of this paper is to facilitate the continuation of research efforts and to encourage bigger participation of researchers in database systems in the field of civil engineering.

Keywords: BIM data, machine learning, civil engineering modelling.

1. Introduction

Building Information Modeling (BIM) is nowadays widely used in architecture, engineering and construction industry (AEC). The building and construction industry employs currently about 7 percent of the world's working-age population and is one of the world economy's largest sectors. It is estimated that about $10 trillion is spent on construction-related goods and services every year. In the last decade, the acceptance and actual use of BIM has increased significantly within the building community. It has largely contributed to the process of eliminating faults in designs. BIM allows architects and engineers to create 3D simulations of the desired structures which contain significantly more information on the actual structures than drawings produced using traditional Computer Aided Drafting CAD systems. As a result, BIM has become more and more present in the construction industry.

BIM technology enables to represent syntactic and semantic building information with respect to the entire life cycle of designed objects, from the design phase, through construction to the facility management phase. BIM includes information about the

---

a

Ślusarczyk, Strug

elements and spaces within buildings, their constituting elements, their interrelations, properties and performance. The project created in BIM technology can be treated as a database that allows to record both technical information about building elements, and its purpose and history. However, although BIM is information rich, not all knowledge is explicitly stated. It seems that machine learning approaches suit well to deduce implicit knowledge from BIM models. Contrary to querying approaches used to extract knowledge from building models[1,2,3],which are tailored to specific scenarios with predefined outcomes, machine learning methods are able to detect patterns and make predictions.

Using machine learning (ML) and artificial intelligence (AI) in AEC industry is a promising research direction. It has to be noted that while both ML and AI are rapidly developing across many other industries, the construction industry is lagging behind in the rate by which improvements are introduced. The usage of BIM could be seen as a tool to revert this trend, but software tools able to implement BIM still require quite laborious routine tasks to properly execute BIM.

The structure of data in BIM, where knowledge is represented in an object oriented way, is ideal for analytical purposes and the application of machine learning techniques[4] . ML is related to extrapolating object behaviours and generating logical responses from information provided by examples, and enabling a computer to gradually learn. Various classification algorithms, anomaly detection, and time series analysis can be used in respect to BIM. Classification algorithms can be used for example to predict characteristics of flats and their sale demand, the likelihood of construction delays, or to diagnose the assets in historic buildings. Anomaly detection is useful in assessment of architectural models, discovering modelling errors, while time series analysis can be applied to make maintenance predictions or renovations planning. In order to make predictions or detect patterns, metrics about buildings are to be specified. They serve as a label to building models which allows for measuring their performance.

This paper is and extended version of the conference paper Ref. 5.. A significant number of new papers have been carefully selected for this extension. The new paper add both deeper analysis of previously presented somains of applications as well as reaching to other applications withi the BIM methodology.

The objective of this paper is to present highlights of references pertaining to machine learning in Building Information Modelling. It complements previously published literature survey articles in order to provide insight into the development of artificial intelligence in BIM, underline the hotspots of current research in this domain, and facilitate continued research efforts. The paper summarizes recently developed theories and methods applied in BIM-based knowledge processing, extraction, and semantic enrichment of BIM models. They include neural networks, decision trees, logistic regression, affinity propagation clustering, term frequency, random forests, SVM, as well as Bayesian networks. The paper provides an overview of the advances of machine learning methods applied in BIM.

2. BIM data - IFC filess

The information about a building created in any software can be exported to different formats. Each commercial application has its own file type to store building data, but all of them can also export building information to an IFC file. The file format IFC has become de facto standard way of interchanging and storing BIM data[6]. It is an interoperable BIM standard for CAD applications, which supports a full range of data exchange among different disciplines and heterogeneous applications. Information retrieved from IFC files can be used by many different applications.
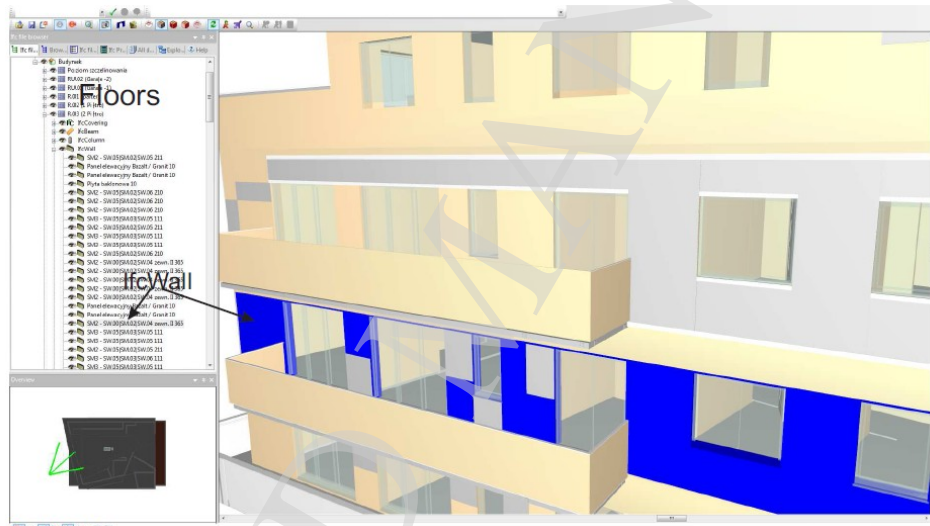

Fig.1 An example of IFC data visualization

IFC specifies different types of building entities and their basic properties. It defines an EXPRESS based entity-relationship model, which consists of several hundred entities organized into an object-based inheritance hierarchy. All the entities in IFC are divided into rooted and non-rooted ones. While the first ones are derived from IfcRoot and have identity (a GUID), attributes for name, description, and revision control, the other ones (non-rooted) do not have identity and their instances exist only if they are referenced from a rooted instance directly or indirectly. IfcRoot is subdivided into three concepts: object definitions, relationships, and property sets:

- IfcObjectDefinition captures tangible object occurrences and types
- IfcRelationship captures relationships among objects
- IfcPropertyDefinition captures dynamically extensible properties of objects

3

Ślusarczyk, Strug

In Fig. 1 a fragment of an example visualization of an IFC file representing a multi-storey building is depicted. In the top left-hand side panel, the hierarchical structure of the file is shown. The floor elements describe the storeys of the building. One of the floor elements is developed to show its different component entities of the type IfcBeam, IfcColumn and IfcWall in a more detail way. The component entities of one of the IfcWall elements are shown and their visualization is depicted by the darker colour in the right-hand side panel in Fig. 1 .
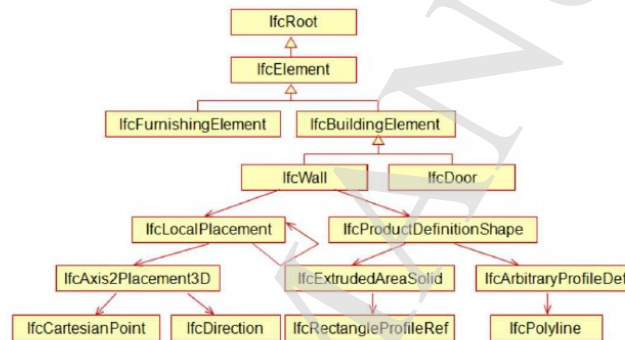

Fig. 2 IfcWall and IfcDoor within IFC hierarchy

The entities that can be used in IFC include building components like IfcWall, IfcDoor, IfcWindow, geometry such as IfcExtrudedAreaSolid, and basic constructs such as IfcCartesianPoint.

The most often used building elements are IfcSpace, IfcDoor, IfcWall, IfcStair and IfcWindow. According to the IFC 2x Edition3 Model Implementation Guide[6] and the IFC specification[6] the above mentioned classes can be described as follows:

1. IfcSpace is the instance used to represent a space as the area or volume of a functional region. It is often associated with the class IfcBuildingStorey representing one floor (the building itself is an "aggregation" of several storeys) or with IfcSite, which represents the construction site. A space in the building is usually associated with certain functions (e.g., kitchen, bathroom). These functions are specified by attributes of the class IfcSpace (Name, LongName, Description).
2. IfcWall is the instance used to represent a vertical element, which is to merge or split the space. In IFC files two representations of a wall can be distinguished. The subclass IfcWallStandardCase of IfcWall is used for all walls that do not change their thickness (the thickness of a wall is the sum of the materials used).

4

IfcWall is used for all other walls, in particular for constructs of varying thickness and for the walls with non-rectangular cross-sections.

3. IfcStair represents a vertical passage allowing for moving from one floor to the other. It can contain an intermediate landing. Instances of IfcStairs are treated as containers, by which we refer to component elements as IfcStairFlight using IfcRelAggregates.

4. IfcDoor represents a building element used to provide access to a specific area or room. Parameters of IfcDoors specify dimensions, an opening direction and a style of the door (IfcDoorStyle). IfcDoor is a subclass of IfcBuildingElement and a superclass of IfcDoorStandardCase. Door instances are usually located in a space IfcOpeningElement to which we refer by IfcRelFillsElement.

5. IfcWindow represents a building element used to fill vertical or near-vertical openings in walls or roofs. It provides a view, light and fresh air. The dimensions of the window and its shape can be found in IfcWindowStyle, to which we refer by IfcRelDefinesByType. IfcWindow is a subclass of IfcBuildingElement and a superclass of IfcWindowStandardCase. Window instances are placed in a space IfcOpeningElement to which we refer by IfcRelFillsElement.

The above mentioned instances inherit from IfcProduct class, which allows for determining their positions using geometrical entities like IfcLocalPlacement and relationships like IfcPlacementRelTo. The place of IfcWall and IfcDoor within the IFC hierarchy is depicted in Fig. 2 (from[8]).

3. Machine learning techniques for BIM

Many recent studies address the need for semantic integrity in BIM models as it allows for sharing these models between project teams, performing domain-specific analyses on BIM model data and checking BIM model deliverables by owners. Moreover, information important from the point of view of many applications lies implicit in the interrelation between building elements. Therefore, several approaches directed at extracting implicit data from building models have been presented. They explore mainly an inductive reasoning approach, which uses machine learning and deep learning, where rules are learned and generalized from the data, to ensure the integrity of BIM element classifications. Bloch and Sacks showed that a machine learning-based approach is more successful in automatically classifying spaces than deductive reasoning, which is based on a rule-inferencing approach[9].

Ślusarczyk, Strug

3.1. *Learning semantic information*

3.1.1. *Space classification*

An unsupervised learning method for mining the IFC based BIM data by exploring interrelations between building spaces is presented in[10]. A method of extracting features which are then used in the affinity propagation clustering algorithm to get spaces with similar usage functions, is proposed. The method allows for automatic learning of functional knowledge from building space structures. The physical properties of each space and their boundary relationships in BIM model are extracted from the IFC file based on BIM data. Then boundary graphs with space boundary relationships, where properties of each space propagate along the edges, are built. Features of building spaces are extracted from the space boundary graphs. Based on these features and the graph representation of the building structure, the adapted affinity propagation algorithm performs building space clustering analysis, in order to get representative samples of building spaces. The experimental results performed on a real-world BIM dataset containing 595 spaces from a 20-storey building show that building spaces with typical usage functions, like senior offices, open offices and circulation spaces, can be discovered by the unsupervised learning algorithm.

In Ref.4 rooms in a housing unit were named according to their use (dining room/lounge, kitchen, bedroom, etc.) based on their geometry. The different types of rooms in BIM are usually labelled entirely by hand by the expert designer. Using ML algorithms to automate this type of task considerably reduces the computational time. Three different classification algorithms, namely decision trees, logistic regressions, and neural networks, were used to solve the problem of labelling rooms according to their function. The input data from which the algorithms are to learn, consist of rooms in the housing unit whose function was labelled by hand previously. The data were obtained from two models of housing developments created with Autodesk Revit, each consisting of more than 200 housing units, from which some rooms were extracted. The algorithms were trained with one project and evaluated with another, which ensures that the results could be extrapolated to other projects. In order to compare the sensitivity of the amount of information available, two different data sets were created for each model. The first model only includes the information obtained directly (or via formulae) from the Revit schedules. The second model also includes information that can only be extracted or calculated by C# programming using the Revit API. Predictions made by logistic regression or neural network with complete data were about 80%–90% accurate in predicting room use. The most common errors are mistaking kitchens for bedrooms and classifying corridors or bedrooms as bathrooms.

Experimental approaches directed at extracting implicit data from building models were also researched[11]. Both unsupervised and supervised machine learning of BIM models is considered. Supervised machine learning approach, which is based on a neural network, is able to classify floor plans according to its intended function. By looking at the spatial configurations of floor plans, a neural network was trained to differentiate

Machine Learning Methods in BIM-based Applications – a Review

between residential and institutional facilities. This approach can be used to complete missing attributes in datasets, where information pertaining to intended function and use is fragmented and incomplete. Both supervised and unsupervised learning algorithms assess a building, by means of a set of its characteristic features. The IFC Machine Learning platform presented in this paper is built on top of the DURAARK IFC metadata extractor[12]. This tool is able to extract literal values, aggregates and derives values from IFC SPF files. In Ref 13 a dataset of IFC files has been proposed to facilitate the comparison of the classification results for different IFC entities. The unsupervised learning approach is used in Ref. 11 to show anomalies in building models. The obtained results make it possible to flag uncommon situations (like unusual large overhangs or an unusual confluence of several building elements) that might need additional checks or coordination, and therefore reduce the failure costs in the construction industry.

In Ref. 9 machine learning algorithms are used to the problem of classification room types in residential apartments. Classification of room types and space labeling are important for the design process, compliance checking, management operations, and many building analysis tasks. The dataset used for supervised machine learning processes contains 32 BIM apartment models. The classification and labeling of room types in this work is based on their function, and assumes that spaces do not have dual function. A multiclass neural network was used with a total of 150 spaces in the dataset for the training process. The dataset was split to 70% for training and 30% for validation of the trained model, which resulted in an 82% correctly classified validation set. The building objects were classified based on the five local features, area, number of doors, number of windows, number of room boundary lines, and floor level offset, and one connecting feature being a direct access. The obtained results showed that machine learning is directly applicable to the space classification problem.

Four different machine learning methods are used to categorize images extracted from BIM of building designs[14]. BIM data are separated into three categories: apartment buildings, industrial building and others. The first method is based on classical machine learning, where Histogram of Oriented Gradients (HOG) is used to extract features, and a Support Vector Machine (SVM) is used for classification. The other three methods are based on deep learning. The first two use pre-trained Convolutional Neural Networks (CNN) (a MobileNet[15] and a Residual Network[16]). The third one is a CNN with a randomly generated structure. A database of 240 images extracted from 60 BIM virtual representations is used to validate the classification precision of the models. The accuracy achieved by the HOG+SVM model is 57%, while for the neural networks it is above 89%. The approach shows that it is possible to automatically categorize a structure type from a BIM representation.

Ślusarczyk, Strug

3.1.2. *Object classification*

In Ref. 17 a deep learning approach was proposed to classify 3D models in BIM environments. A ray-based feature extraction algorithm was used to extract the features of 3D models and form a feature matrix. This matrix is applied by the deep belief network (DBN) constructed by restricted Boltzmann machines, which classifies the models. The database included 1,814 models, covering most types of models common in daily life, which were divided into a training set and a testing set, each with 907 models. The results show that the proposed approach achieves accuracy of 85.5%.

Koo et al. applied the SVM technique to classify building elements for checking the semantic integrity of mappings between BIM elements and IFC classes[18]. The SVM was trained to distinguish model elements from a dataset of 4187 unique elements collected from six architectural BIM models, based on their geometric and relational features. In the first stage, the SVM was tested to classify the elements with respect to eight IFC classes (wall, slab, beam, door, window, railing, covering, and column). The dataset was divided into training and test sets by a ratio of eight to two resulting in a test set of 838 elements on which 10-fold cross validation was performed. In the second stage, the SVM was further tested to distinguish between the element subtypes within three IFC classes, IfcDoor (single and double door), IfcWindow (single and double window), and IfcColumn (rectangular and non-rectangular shaped column). The same training to test ratio was used and 10-fold cross validation was performed to optimize hyperparameters. The approach showed 94.39% accuracy of classifying building elements in the first stage, while 100% for the window and column types, and 96% for door types in the second stage.

As it is difficult to ensure the BIM data integrity due to errors, mistakes or omissions of object information, Kim et al. used 2D CNN to recognize and classify unknown objects as BIM instances[19]. The presented approach classifies basic architectural BIM elements, as well as furniture entities according to their images using a 3D object recognition technique. In the first stage the class of an object is recognized (slab, wall, ceiling, column, staircase, door, seating furniture and toilet fixture), while in the second one, subtypes of door, seating furniture and toilet fixture are covered. The dataset consists of 820 objects spanning 10 classes. The model input consists of multi-view 2D rendering images of BIM objects. The training of the recognition models is processed by fine-tuning the Inception-v3 model[20] that trains the ImageNet[21] dataset on TensorFlow framework[22]. 80% of the data is used for training and 20% for validation. The accuracy of training for the classification of building elements is 96%. The second stage training results showed the average of 96.7% accuracy of classification of sub-types of doors, seating and toilet fixture.

3D geometric deep neural networks were used for automatic classification of BIM elements by Koo et al.[23]. Multi-view CNN (MVCNN)[24] and PointNet[25] were investigated to determine their applicability in extracting unique features of door (IfcDoor) and wall (IfcWall) element subtypes. MVCNN builds classifiers of 3D shapes from 2D image renderings of these shapes, while PointNet is a deep neural network, which takes point

Machine Learning Methods in BIM-based Applications – a Review

clouds directly as input to classify and segment 3D artifacts. Firstly 478 door elements and 364 wall elements were collected and divided into their respective subtypes as single, double, sliding and revolving door, and a generic wall, wall with window opening, and wall with door opening. The samples were divided using a 7:3 ratio to train and test the learning models. Test results indicated MVCNN as having the best prediction performance with accuracy of 92% and 95% for door and wall element subtypes, respectively, while PointNet's accuracy was hampered by resolution loss due to selective use of point cloud data.

Geometric deep learning models were also used to classify infrastructure BIM elements for checking the integrity of element-to-entity mappings in BIM models[26]. MVCNN and PointNet models were trained and tested to classify 10 types of commonly used BIM elements in road infrastructure. They include columns, three types of culverts, four types of retaining walls, sumps and wing walls. A dataset consists of 1496 3D elements, which were divided into training and test sets, using a 7:3 ratio. Results revealed MVCNN as the superior model with accuracy 98%, compared with the PointNet's corresponding value of 83%.

Collins et al. encode BIM objects using two kinds of graph encodings and utilize a graph convolutional network to perform classification based on the 3D shape only[27]. A Graph Convolutional Network (GCN) uses the native triangle-mesh and automatically creates meaningful local features for subsequent classification. The data are extracted from 22 IFC tagged files and divided into training and test sets, using an 8:2 ratio. The dataset consists of structural elements (IfcWall, IfcSlab, IfcColumn, IfcWindow, IfcDoor, IfcStair IfcRailing), the equipment (IfcFlowTerminal, IfcFlowSegment, IfcFlowFitting, IfcDistributionControlElement, IfcFlowController) and the interior furniture (IfcFurnishingElements). The method reaches an accuracy of up to 85% on the dataset collected by authors.

Luo et al. propose a geometric-relational deep learning framework for BIM object classification[28], which boosts the classification methods based on object geometry with relational information. The proposed framework consists of three main modules, i.e., a geometric feature extractor, relational feature extractor and feature fusion module. The geometric feature extractor learns to represent a BIM object's shape as a geometric descriptor. The relational feature extractor extracts relational features of different levels and connects them to form the relational descriptor. The feature fusion module mixes the two descriptors and outputs the final object descriptor. Three proposed relational models RMVCNN RDGCNN RMVViT are based on MVCNN, DGCNN[29] and MVViT (a 3D deep learning model adapted from Vision Transformer[30]), respectively, which act as geometric feature extractors. The IFCNet++ dataset, where selected representative relational features are attached to each BIM object, was proposed and used in all the experiments for training and testing the described models. The dataset contains 9228

Ślusarczyk, Strug

objects belonging to 11 types. Four subtypes of IfcRelationship that appear in most IFC files (IfcRelConnectsElement, IfcRelFillsElement, IfcRelAggregates and IfcRelVoidsElement) were selected. The data were divided into training and test sets using a 7:3 ratio. The results show that RMVCNN reaches an accuracy higher than 98% on 9 of the 11 types, which means that it can effectively learn geometric and relational features of most types. Accuracy on the other two models is slightly lower than 90%. Experiments show that relational features are representative enough to help the geometric methods achieve a performance gain.

3.2. *Semantic enrichment of BIM models from point clouds*

The other group of papers considers the use of machine learning algorithms for semantic enrichment of BIM models obtained from point cloud data[31]. Semantic enrichment techniques solve the interoperability problem by exploiting existing numeric, geometric, or relational information in the model to infer new semantic information. In this way the time-consuming process of manually creating 3D models useful for architectural and civil engineering applications can be avoided. Semantic enrichment encompasses classification of building objects, aggregation and grouping of building elements, implementing associations to reflect connections and numbering[32], unique identification, completion of missing objects, and reconstruction of occluded objects. Then the classification of the model as a whole, or of particular assemblies or objects within the model in respect to code compliance, can be performed. BIM objects derive many of their properties from their class, making object classification crucial for reuse in different analysis tasks, like spatial validation of a BIM model, quantity take-off and cost estimation.

Xiong et al. use machine learning for classifying and labeling surfaces obtained from a laser scan, in order to semantically enrich a BIM model[33]. They use both the shape features, which are referred to as local features, and the spatial relationships (orthogonal, parallel, adjacent and coplanar) referred to as contextual features. A context-based machine learning algorithm, called stacked learning[34], is used to label patches extracted from a voxelized version of the input point cloud. The main constructive objects of a building, i.e., walls, ceilings and floors, were classified, and separated from other objects obtained from the scan that are considered clutter. The method achieved an average 85% accuracy over 4 classes. Then the SVM algorithm is used to encode the characteristics of an opening shape and location, which allows the algorithm to estimate the shape of window and doorway openings even when they are partially occluded. The method was evaluated on a large, highly cluttered data set of a two-storey schoolhouse building containing 40 rooms. The facility was scanned from 225 locations resulting in over 3 billion 3D measurements. The SVM algorithm was able to detect window and door openings with 88% accuracy.

In Ref. 35 Emunds et al. propose an efficient neural network to learn from point cloud representation of BIM objects. Sparse Residual CNN for Semantic Enrichment of BIM

Machine Learning Methods in BIM-based Applications – a Review

models (SpaRSE-BIM), a neural network model based on sparse convolutions, is used for the classification of IFC-based geometry and semantic enrichment of BIM models. Sparse convolutions only store the non-empty parts of the space and compute outputs for predefined coordinates, making them more efficient and faster than regular 3D convolutions on sparse data. The structure of SpaRSE-BIM is made up of three building blocks, the Sparse Conv Block, the MLP Block, and the Sparse Residual Block, which is the main building block of the network. SpaRSE-BIM processes point clouds. The experiments were performed on the IFCNetCore and BIMGEOM datasets. IFCNet provides a dataset of 19,613 individual IFC objects from 65 IFC classes, while its subset, called IFCNetCore, provides 7,930 objects from 20 classes. BIMGEOM consists of 10,146 objects of structural elements, technical equipment and interior furniture and contains 13 IFC classes. The results indicate that SpaRSE-BIM achieves overall accuracy of 81.59% on IFCNetCore and 90.47% on BIMGEOM.

The machine learning methods for both semantic enrichment and automating design review using BIM models is proposed[36]. The approach was applied to identify security rooms and their walls within the BIM model created in the design process. A two-class decision forest classification algorithm was chosen for this experiment[37]. It was implemented on a data set of models with 642 security rooms with non-regular complex geometry, arranged in 64 shafts. The dataset for training contained 448 spaces, 278 of which were security rooms compliant to the described code clause and the remaining 170 were other rooms or open spaces. The spaces were organized in 64 vertical shafts, each of which was comprised of security rooms and other spaces. Running a 10-fold cross validation algorithm achieved 88% accuracy. Now the authors explore the possibility of developing a deep neural network to classify the rooms when the only input data are a wall schedule, a room schedule and a table of the relationships between them.

3.3. *Supporting building design*

Different types of data available in BIM models can be used to support the building design phase in the AEC industry. They help to optimize the design, detect clashes, determine the building energy consumption in order to improve sustainability of the construction, better understand client's needs, and make better informed decisions.

Chen et al. proposed a framework combining BIM with the least square support vector machine (LSSVM) and non-dominated sorting genetic algorithm-II (NSGA-II) to study the influence of building envelope parameters on building energy consumption[38]. The LSSVM is utilized to learn the data set to establish a prediction model between the building envelope and energy consumption, which is used as the fitness function of the building energy consumption in NSGA-II. An NSGA-II multi-objective optimization model with the objective of minimizing building energy consumption and maximizing

Ślusarczyk, Strug

indoor thermal comfort is established. The proposed framework was used to optimize the design of a school building. The results show that the LSSVM has great ability to predict building energy consumption, achieving a prediction accuracy of 95.49%. The developed LSSVM-NSGA II hybrid approach can effectively improve the building's energy consumption and thermal comfort with the building energy consumption reduced by 10.6% and the thermal comfort increased by 32.2% compared to the initial values, respectively. A unified framework for the extraction and integration of BIM models and building energy consumption monitoring data has been proposed[39]. The energy usage anomaly diagnosis methods have been devised using integrated static and dynamic data, which employ density-based clustering and artificial neural network regression using the integrated data. The multi-layered backpropagation neural network model regressing over input data from a multitude of buildings, with building static data treated as input parameters, is used to identify implicit anomalies that are beyond the scope of clustering methods.

In Ref. 40 machine learning is used to predict building thermal energy performance. A component-based approach that develops machine learning models for a parameterized whole building design and for parameterized components of the design is described. ML components have been developed at two levels of decomposition, namely construction-level components (wall, window, floor, roof, etc.) with a static prediction of total yearly heating and cooling energy consumption as response, and zone-level components with dynamic monthly consumption as a response. The development of ML components representing construction elements uses neural networks with one hidden layer per component. Training data were split into a training set (85%) and a set used for independent cross validation of components after testing (15%). Two algorithms were used for training, namely the Levenberg-Marquardt algorithm and Bayesian regularization. Results in test cases show that high prediction quality is achieved with errors of 3.7% for cooling and 3.9% for heating. In Ref. 41 the component-based ML using artificial neural networks is used for energy performance prediction in early stages of building design. As the application of BIM in early design phases requires the support of different levels of detail, the approach links multiLOD building models with ML components. This approach was illustrated and validated by a test case with a medium-sized office building. For validation purposes, the results of the ML predictions for 60 different design configurations were considered, showing that the average error was no higher than 3.5% for cooling and 3.8% for heating.

In Ref. 42 a system based on BIM and ANN, which evaluates indoor personal thermal comfort and energy of the interior space design, is proposed. It integrates thermal information and a building model. The back propagation ANN predictive model of thermal comfort takes three environment parameters (air temperature, air humidity, and wind speed around the person), three human state parameters (human metabolism rate, clothing thermal resistance, and the body position) and four body parameters (gender, age, height, and weight) as inputs. The proposed plugin of BIM provides thermal comfort evaluation results and energy saving optimization suggestions for interior space design by simulating the energy consumption index of different design schemes. In the data test, the

Machine Learning Methods in BIM-based Applications – a Review

Mean Squared Error of the established ANN model was equal to 0.39. In Ref. 43 a framework based on BIM and machine learning data-driven models, with a goal to analyze the optimum thermal comfort for indoor environments with the effect of natural ventilation, is proposed. Sufficient natural ventilation in buildings is important for reducing the energy consumption of mechanical ventilation while maintaining a comfortable indoor environment for occupants. BIM provides geometrical and semantic information of the built environment, which are needed for specifying boundary conditions of computational fluid dynamics (CFD) simulation. This simulation is conducted to obtain the flow field and temperature distribution, the results of which determine the thermal comfort index in a ventilated environment. BIM-CFD provides spatial data, boundary conditions, indoor environmental parameters, and the thermal comfort index for machine learning to construct robust data-driven models to empower the predictive analysis. In the neural network, the spatial features such as zone adjacency and connectivity, and the impact of airflow between zones are represented in the form of the adjacency matrix.

A system which aims to assist architects in designing buildings with minimal energy costs while meeting the functional purpose of the designed facility, is proposed in Ref. 44. It integrates BIM and ANN to predict energy cost of residential buildings at the conceptual stage of architectural design. BIM enables the designer to produce different scenarios of architectural design, while the ANN model is used to predict the energy cost of the designs defined in BIM. Data sets used for developing the prediction model contain factors influencing the cost of energy such as the age of the existing residential building, location, building area, number of occupants, glazing type, insulation, a building envelope system and air conditioner type. 186 data sets were used to build the model, 144 data sets were utilised for training, 16 data sets were used testing the model and 26 data sets were used for validation. The results show that the developed ANN-based model provides accuracy of 78% in estimating the energy cost of the residential buildings. A method which integrates Building Information Modelling with artificial neural network model for limiting the deviation between predicted and actual energy consumption rates is proposed in Ref. 45. This deviation is due to the fact that energy performance tools, which are used to simulate the expected energy consumption of a given building do not take into account the occupants' behaviour. In this study optimum occupant behaviour profiles that are generated via the ANNs are inputted into the BIM file in order to run necessary energy simulations to validate the estimates of energy use of the building system. The results show that the heating and cooling energy consumption can be predicted with an average of 10.4% deviation from the actual performance measured in the buildings.

Automatically detecting clashes in designs by merging different BIM models supports the building design stage. In Ref. 46 six kinds of supervised machine learning

Ślusarczyk, Strug

algorithms (decision tree, random forest, Jrip-based rule methods, binary logistic regression, naïve Bayes, and Bayesian network) are used to automatically distinguish relevant and irrelevant clashes. Ref. 47 develops a method that automatically filters out irrelevant clashes from the BIM-generated clash detection reports by combining the two techniques of rule-based reasoning and supervised machine learning. BIM software is used for clash detection and rule-based reasoning is used to preliminarily determine the type of clashes. The results of this classification are added into the dataset of machine learning for training and testing classifiers (k-NN, SVM, DT, random forest, bagging). The average predictive performance accuracy obtained by using the hybrid method is up to 96%. In Ref. 48 an artificial intelligence system for design clash resolution was developed with machine learning and heuristic optimizing techniques. A parallel back-propagation neural network was employed as a learning procedure to extract the clash resolution knowledge patterns. Based on these patterns, simulated annealing was adopted to develop a system for automatically resolving design clashes from the constructor's perspective. To achieve this in BIM models, the proposed system utilizes the API provided by the BIM software as the portal to detect clashes, retrieve the geometric and spatial profiles of elements, and revise the elements.

In Ref. 49 a natural language processing and deep learning-based approach for supporting a BIM-based automated rule checking system is described. The semantic analysis for building regulations is performed in two steps: first the meaning of words and the topic of sentences are learned by the neural net-based word embedding technique (word2vec model) and then sentence classification is conducted with a deep learning model.

3.4. *Building condition diagnosis*

Semantically enriched BIM models are often used in the heritage industry to manage, analyse and diagnose the assets at varying stages of the conservation process. Authors of Ref. 50 uses machine learning techniques to automatically classify heritage buildings. SVM are proposed to extract the main structural components such as floors, ceilings, roofs, walls and beams. The proposed semantic labelling of the objects is based on features, which encode both local (surface area, orientation, dimensions) and contextual information (normal similarity, coplanarity, paralellity, proximity, topology) are extracted form training data sets. The proposed automated feature extraction algorithm combined with an SVM classifier takes the preprocessed data in the form of planar triangular meshes and outputs the classified objects. The algorithm was trained and tested using real data of a variety of existing buildings, like houses, offices, industrial buildings and churches. 10 structures representing different types of buildings were selected for the evaluation. The average accuracy of the model is 81%. The experiments prove that the approach reliably labels entire point cloud data sets and can effectively support experts in documenting and processing heritage assets.

Machine learning methods are also used for defect classification in masonry walls of historic buildings[51]. First, the process of Scan-to-BIM, which automatically segments point clouds of ashlar masonry walls into their constitutive elements, is presented. Then the machine learning based approach to classification of common types of wall defects, that considers both the geometry and colour information of the acquired point clouds, is described. The found defects are recorded in a structured manner within the BIM model, which allows for monitoring the effects of deterioration. A supervised logistic regression algorithm has been employed to classify different types of decay using parameters of roughness of stones and dispersion of colour in stones. Stones labelled as 'defective' by experts are used for training the classifier, which is subsequently employed to label new data. The proposed approach has been tested on data from the main façade of the Royal Chapel in Stirling Castle, Scotland. For the training process samples of three classes of decay (erosion, mechanical damage and discolouration) were used. 15 samples (5 of each class) were included in the test set, obtaining a global accuracy of 93.3% in the classification.

3.5. *BIM enhancement in the facility management context*

Semantic enrichment by integrating facility management (FM) information with a Building Information Model is presented in Ref. 52. At first, various machine learning algorithms, which analyse the unstructured text from occupant-generated work orders and classify it by category and subcategory with high accuracy, have been investigated. Then, three learning methods, Term Frequency (TF), Term Frequency-Inverse Document Frequency (TF-IDF), and Random Forest classifier, were applied to perform this classification. A set of 155.00 historical WOs was used for model development and testing textual classification. Classifier prediction accuracies ranged from 46.6% to 81.3% for classification by detailed subcategory. It increased to 68% for simple TF and to 90% for Random Forest when the dataset included only the ten most common subcategories. An FM-BIM integration provides FM teams with spatio-temporal visualization of the work order categories across a series of buildings and helps prioritize maintenance tasks. The paper shows that machine learning can be applied to support FM activities and can enhance BIM use in the FM context.

To provide a good maintenance strategy for building facilities Cheng et al. developed a data-driven predictive maintenance planning framework based on BIM and IoT technologies[53]. Data collection and data integration among the BIM models, FM system, and IoT network are undertaken in the information layer, while ANN and SVM, are used in the application layer to predict the future condition of mechanical, electrical and plumbing (MEP) components. BIM models provide accurate and complete information for predictive maintenance, including component type, dimensions, materials, capacity,

Ślusarczyk, Strug

location, and installation year. An IoT sensor network was established to gather sensor data from building facilities and environment during the operation period. The prediction models are trained based on the continuously updated real-time sensor data and the accumulated maintenance record. The input datasets are randomly divided into three sets: 80% for model training, 10% for cross validation, and 10% for model testing. The accuracy of prediction of MEP components condition is equal to 96.547% for SVM and 96.422% for ANN. The results show that the constantly updated data obtained from the information layer together with the machine learning algorithms in the application layer can efficiently predict the future condition of MEP components, prevent the failure and extend the lifetime of these components. In Ref. 54 a decision support system for assessing the quality of a large building stock is proposed. The system is meant to help owners of large building stocks in managing a huge number and variety of buildings, taking into consideration a lot of aspects related to different technical issues, such as accessibility, energy efficiency, life safety, fire protection and seismic vulnerability. It can extract relevant information from a BIM database and evaluate the compliance of existing buildings with a set of rules by means of the application of Bayesian Networks.

A knowledge-based intelligent Building Energy Management System is considered in Ref. 55 . To support the integration of multiple data sources and improve interoperability, the solution makes use of BIM principles and semantic web technologies in the form of a holistic knowledge base, into which the rules are integrated. The developed ontology is based upon the Industry Foundation Classes (IFC). The populated rules are a mix of an intelligent rule generation approach using Artificial Neural Network and Genetic Algorithms, and also data mining rules using Decision Tree techniques on historical data. The rules are queried by the fuzzy reasoner together with the currently monitored state of the building to generate suggestions for the FM.

3.6. *Knowledge extraction from BIM*

Methods of extracting knowledge from existing BIM models with the use of machine learning are also considered. In Ref. 2 the notion of the building fingerprint is used to capture the main characteristics of a building design. It serves as a measure of similarity which allows for finding a suitable reference for a given problem. The fingerprint is based on accessibility and adjacency relationships among spaces within a building model. Therefore, the authors retrieve accessibility and adjacency relationships among spaces encoded in IFC models, and build accessibility and adjacency graphs. Building fingerprints are automatically generated based on a spatial semantic query language for BIM and applied as indexes of the building model repository.

In Ref. 56 the ensembles of artificial neural networks are used for cost estimation based on building information models. Building information model serves as a repository of data. A regression model for cost estimation of building floor structural frames, as higher-level elements, was created. This approach was continued in Ref. 57 , where a model based on the support vector regression and radial basis kernel functions was developed and proposed to predict construction costs of building floor structural frames.

Machine Learning Methods in BIM-based Applications – a Review

The ability to extract certain information about the building and structural members of the floor frames from BIM was combined with the capabilities of machine learning. The training data for machine learning included inputs that represented features of the building and structural members' belonging to the floor structural frames and outputs that represented corresponding real life cost estimates of the floor structural frames. The obtained results show that the proposed model allows predicting costs with satisfactory accuracy.

A BIM and ML integration framework for automated property valuation was proposed[58]. It contains a fundamental database interpretation, an IFC-based information extraction and an automated valuation model based on genetic algorithm optimized machine learning (the gradient boosting regression- GA-GBR). The value-relevant attributes or properties for property valuation were achieved through quantitative analysis of 174 archival documents. The required value-specific design information was extracted automatically from an IFC-based BIM instance model to support property valuation. The GA-GBR model was used to evaluate property data in American real estate market. The experimental dataset was divided into 20 groups with 1000 houses data in each of them and randomly split into training set (70%) and testing set (30%). The 56 input features were further used for problem encoding in the GA-GBR training section. The results indicate that GA-GBR model achieves accuracy of 89.4%.

In Ref. 59 a Bayesian network is used to gain knowledge about existing bridges by the use of data in the bridge management system in order to support bridge engineering design. Two variants for the generation of the Bayesian network were considered: the Tree Augmented Naive Bayesian Network Algorithm and the manual creation of the network by a knowledge engineer. Both variants permit the determination of design bridge parameters on the basis of given boundary conditions. The first variant fits better to the generated training data sets, while the manually created network supports the bridge design process in a more intuitive way. The extracted knowledge is used for new bridge design problems to generate BIM models for possible bridge variants.

4. Discussion

In this paper a significant number of recent paper related to the application of different AI methods in BIM has been prresnted. As the papaers have been carefully selected by the authors rather then obtain via a keyword search the present state-of-the art results. Theere are several important obsarvations as well as lessens that can be drawn from this analysis. The most important one is that the majority of the best results are obtained not by using one methos but rather by combining different approaches. The methods most often used

Ślusarczyk, Strug

are summarized in the table below. It can be observed that different variants of neural network and convolution neural networks are the method of choice for the majority of the researchers. Neverthe less other methods like Bayesyan classifier, decision trees or genetic algorithms wers also used with success either on its own or, more often, in combination with other methods,

Method	Domain of application	Papers
Neural networks	Semantic classification supporting design facility management knowledge extraction	4,9,10,11 39,40,41,42,43,44,45,48 53,55 56
Convolution neural networks	Semantic classification semantic enrichment	14,19,23,26,28 35,
Graph neural networks	semantic classification	27
Decision trees/forests	Semantic classification semantic enrichment supporting design facility management	11 36 46,47 52
SVM	Semantic classification semantic enrichment supporting design building condition facility management knowledge extraction	14,18. 33 38 50 53 57
Deep belief networks	3D model classification	17
Bayesian networks	supporting design facility management knowledge extraction	46 54 59
Genetic algorithms	supporting design facility management knowledge extraction	38 55 58
Rule based methods	supporting design	46,47

Table 1. Methods of Ai used for different domains of applications in BIM.

Machine Learning Methods in BIM-based Applications – a Review

5. Conclusions

Traditional methods for modeling and optimizing complex structure systems require huge amounts of computing resources. Artificial-intelligence-based methods can often provide valuable alternatives for efficiently solving problems in architectural and engineering design, construction and manufacturing.

Machine learning techniques have considerable potential in the development of BIM. The application of classification algorithms would enable machines to do tasks usually done by hand. Results from machine learning on architectural datasets provide a relevant alternative view to explicit querying mechanisms and provide useful insights for more informed decisions in the design and management of buildings. Machine learning might facilitate less experienced users to query complex BIM datasets for project specific insights.

It was shown that machine learning algorithms can learn the key features of a building belonging to a certain category, and this acquired knowledge could be used in the future when designing the methods to automatically design other structure based on BIM historical data. The described techniques can be applied for retrieval, reference, and evaluation of designing, as well as generative design. The presence of historical data combined with the acquired knowledge of the building type key features could help in developing methods for automatically designing building structures with required characteristics. The presented methods can be extended to further subdivide the BIM main categories into sub-categories that could represent different areas of interest in these structures.

Many problems in architectural and engineering design, construction management, and program decision-making, are influenced by many uncertainties, incomplete and imprecise knowledge. It seems that machine learning techniques should be able to fill knowledge gaps in the knowledge bases and therefore they have broad application prospects in the practice of design, construction, manufacturing and management. Object classification integrity is a fundamental requirement that needs to be satisfied in order to ensure effectiveness of using BIM applications. Successful semantic enrichment tools would infer any missing information required by the receiving application, thus alleviating the need for the domain expert to preprocess the building model. They can help inexperienced users to solve complex problems, and can also help experienced users to improve their work efficiency, and share experience.

The authors believe that the widespread use of machine learning methods in BIM community would require a more probabilistic and less deterministic approach to the parameters in the models, and/or the implementation of correction and post-process measures. Such measures could include automatic revision based on criteria established using traditional programming, to pinpoint the elements where ML algorithms make

19

Ślusarczyk, Strug

mistakes. In order to obtain consistent results with sufficient predictive potential, not only the choice of the right ML algorithms is important, but also the choice, quantity and quality of the data used to train the algorithms.

References

1. W.Mazairac, J.Beetz., Bimql? an open query language for building information models,. Adv. Eng. Inform, 27:444–456, 2013.
2. S. Daum, A.Bormann., Automated generation of building fingerprints using a spatio-semantic query language for building information models., In eWork and eBusiness in Architecture, E gineering and Construction: ECPPM 2014} 2014.
3. C.Zhang, J.Beetz, M.Weise., Interoperable validation for ifc building models using open standards, In Special Issue ECPPM 2014 - 10th European Conference on Product and Process Modelling,, volume~20, pages 24--39., 2014.
4. P. E. Núnez-Calzado, I.J. Alarcón-López, D. C. Martínez-Gómez., Machine learning in bim. In EUBIM 2018, 2018.
5. G. Ślusarczyk, B. Strug Machine Learning Methods for BIM Data*ACIIDS (2) 2022: 230-240*

6. BuildingSMART. http://www.buildingsmart-tech.org/,, 2013

7. T. Liebich. Ifc 2x edition3 model implementation guide, https://standards.buildingsmart. org/, 2009.
8. S.Wang, G.Wainer, V.Rajus, R. Woodbury. Occupancy analysis using building information modeling and cell-devs simulation. Vol. 45, 04 2013.
9. T. Bloch, R. Sacks, Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. Automation in Construction. 91. 256-272. 10.1016/j.autcon.2018.03.018. , 2018
10. Ch. Jin, M. Xu, L. Lin, X.4 Zhou., Exploring bim data by graph-based unsupervised learning. Pages 582--589, 01 2018.
11. Th. Krijnen, M. Tamke. Assessing implicit knowledge in bim models with machine learning. Pages 397–406, 01 2015
12. J. Beetz, I. Blmel, S. Dietze, B. Fetahui, U. Gadiraju, M. Hecher, T. Krijnen, M. Lindlar, M. Tamke, R. Wessel, and R. Yu. Duraark: Enrichment and preservation of architectural knowledge. In *How to manage data and knowledge related to interpretative digital 3D reconstructions of Cultural Heritage?,*, LNCS. Springer, 2015.
13. C. Emunds, N. Pauen, V. Richter, J. Frisch and C. van Treeck, IFCNet: A Benchmark Dataset for IFC Entity Classification,Proceedings of the EG-ICE 2021,pp.166–175, Universitaetsverlag der TU Berlin, Berlin, 2021
14. F. Lomio, R. Farinha, M. Laasonen, H. Huttunen. Classification of building information model (bim) structures with deep learning. *arXiv:1808.00601 [cs.CV]*, 2018.
15. A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. 04 2017.

Machine Learning Methods in BIM-based Applications – a Review

16. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition,. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),*, pp. 770–778, 2016.

17. L. Wang, Z.K. Zhao, N. Xu, Deep belief network based 3D models classification in building information modeling, Int. J. Online Eng. 11 (5) (2015) 57–63.

18. Koo, B., La, S., Cho, N.W., Yu, Y.: Using support vector machines to classify building elements for checking the semantic integrity of building information models. Automation in Construction 98, 183–194 (2019)

19. Kim, J., Song, J., Lee, J.K.: Recognizing and classifying unknown object in bim using 2d cnn. In: International Conference on Computer-Aided Architectural Design Futures. pp. 47–57. Springer (2019).

20. Szegedy, C., et al.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016).

21. Deng, J., et al.: Imagenet: a large-scale hierarchical image database. In: Proceedings of Computer Vision and Pattern Recognition, pp. 248–255 (2009).

22. Abadi, M., et al.: Tensorflow: large-scale machine learning on heterogeneous distributed systems, arXiv preprint arXiv:1603.04467 (2016).

23. Koo, B., Jung, R., Yu, Y.: Automatic classification of wall and door bim element subtypes using 3d geometric deep neural networks. Advanced Engineering Informatics 47, 101200 (2021).

24. H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller, Multi-view convolutional neural networks for 3d shape recognition, in: International Conference on Computer Vision Santiago, Chile, 2015, https://doi.org/10.1109/ICCV.2015.114.

25. C. Qi, H. Su, K. Mo, L.J. Guibas, Pointnet: Deep learning on point sets for 3d classification and segmentation, Comput. Vis. Pattern Recogn.1(2) , https://doi.org/10.1109/cvpr.2017.16.

26. Koo, B., Jung, R., Yu, Y., Kim, I.: A geometric deep learning approach for checking element-to-entity mappings in infrastructure building information models. Journal of Computational Design and Engineering 8(1), 239–250 (2021).

27. Collins, F.C., Braun, A., Ringsquandl, M., Hall, D.M., Borrmann, A.: Assessing ifc classes with means of geometric deep learning on different graph encodings. In: Proc. of the 2021 European Conference on Computing in Construction (2021).

28. Luo, H., Gao, G., Huang, H., Ke, Z., Peng, C., & Gu, M. (2022). A Geometric-Relational Deep Learning Framework for BIM Object Classification. *arXiv preprint arXiv:2212.00942*.

29. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph cnn for learning on point clouds. Acm Transactions On Graphics (tog) 38(5), 1–12 (2019).

30. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020).

31. R. Sacks, L. Ma, R. Yosef, A. Borrmann, S. Daum, U. Kattel. Semantic enrichment for building information modeling: procedure for compiling inference rules and operators for

Ślusarczyk, Strug

complex geometry,. *Journal of Computing in Civil Engineering*, 31(6), 2017.

32. L Ma, R. Sacks, U. Kattel, and T. Bloch. 3d object classification using geometric features and pairwise relationships: 3d object classification using features and relationships. *Computer-Aided Civil and Infrastructure Engineering*, 33(2):152164, 2018.

33. A. Xiong, B. Adan, and and D. Huber Akinci. Automatic creation of semantically rich 3d building models from laser scanner data. *Autom. Constr*, 31:325–33, 2013.

34. D. Wolpert. Stacked generalization. *NeuralNetworks*, 5:241 –259, 1992

35. Emunds, C., Pauen, N., Richter, V., Frisch, J., van Treeck, C.: Sparse-bim: Classification of ifc-based geometry via sparse convolutional neural networks. Advanced Engineering Informatics 53, 101641 (2022).

36. Rafael Sacks, Tanya Bloch, Meir Katz, and Raz Yosef. Automating design review with artificial intelligence and bim: State of the art and research framework. In *ASCE International Conference on Computing in Civil Engineering 2019*, 2019.

37. S. B. Kotsiantis, I. Zaharakis, and P. Pintelas. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering,*, 160:3–24, 2007.

38. Bin Chen, Qiong Liu, Hongyu Chen, Lei Wang, Tingting Deng, Limao Zhang, Xianguo Wu, Multiobjective optimization of building energy consumption based on bim-db and lssvm-nsga-ii, J. Cleaner Prod. 294 (2021) 126153, https://doi. org/10.1016/j.jclepro.2021.126153.

39. S. Yuan, Z.-Z. Hu, J.-R. Lin, Y.-Y. Zhang, A framework for the automatic integration and diagnosis of building energy consumption data, Sensors (Switzerland) 21 (4) (2021) 1–19.

40. [P. Geyer and S. Singaravel, Component-based machine learning for performance prediction in building design, Appl. Energy, vol. 228, no. October 2017, pp. 1439–1453, 2018.

41. P. Geyer, M.M. Singh, S. Singaravel, Component-based machine learning for energy performance prediction by multiLOD models in the early phases of building design, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) vol. 10863 LNCS (2018) 516–534.

42. G. Ma, Y. Liu, S. Shang, A building information model (BIM) and artificial neural network (ANN) based system for personal thermal comfort evaluation and energy efficient design of interior space, Sustain. (Switzerland) 11 (2019) Sept.

43. V. Gan, H. Luo, Y. Tan, M. Deng, and H. Kwok, Bim and data-driven predictive analysis of optimum thermal comfort for indoor environment, *Sensors*, vol. 21, no. 13, 2021.

44. A. Alshibani, O.S. Alshamrani, ANN/BIM-based model for predicting the energy cost of residential buildings in Saudi Arabia, J. Taibah Univ. Sci. 11 (6) (2017) 1317–1329.

45. A. W. Hammad, Minimising the deviation between predicted and actual building performance via use of neural networks and BIM, *Buildings*, vol. 9, no. 5, 2019.

46. Yuqing Hu, Daniel Castro-Lacouture, Clash Relevance Prediction Based on Machine Learning, J. Comput. Civil Eng. 33 (2) (2019) 04018060, https://doi. org/10.1061/(ASCE)CP.1943-5487.0000810.

47. Will Y. Lin, Ying-Hua Huang, Filtering of irrelevant clashes detected by BIM software using a hybrid method of rule-based reasoning and supervised machine learning, Appl. Sci. (Switzerland) 9 (24) (2019) 5324, https://doi.org/10.3390/ app9245324.

22

Machine Learning Methods in BIM-based Applications – a Review

48. H. C. Hsu, S. Chang, C. C. Chen, and I. C. Wu, Knowledge-based system for resolving design clashes in building information models, *Automation in Construction*, vol. 110, no. October 2019, p. 103001, 2020.

49. J. Song, J. Kim, J.K. Lee, NLP and deep learning-based analysis of building regulations to support automated rule checking system, International Association for Automation and Robotics in Construction I.A.A.R.C), 2018.

50. M.~Bassier, M.~Vergauwen, and B.~Van Genechten. Automated classification of heritage buildings for as-built bim using machine learning techniques. In {\em ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences}, vol. IV-2/W2, 2017.

51. E.~Valero, A.~Forster, F.~Bosché, C.~Renier, E.~Hyslop, L.~Wilson. High level-of-detail bim and machine learning for automated masonry wall defect surveying. In Proc. of 35th International Symposium on Automation and Robotics in Construction (ISARC 2018)}, 2018.

52. J. McArthur, N. Shahbazi, R. Fok, Ch. Raghubar, B. Bortoluzzi, A. An. Machine learning and bim visualization for maintenance issue classification and enhanced data collection. *Advanced Engineering Informatics*, 38:101–112, 10 2018.

53. Jack C.P. Cheng, Weiwei Chen, Keyu Chen, Qian Wang, Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms, Autom. Constr. 112 (2020) 103087, https:// doi.org/10.1016/j.autcon.2020.103087.

54. A. Corneli, S. Meschini, V. Villa, G.M. Di Giuda, A. Carbonari, A decision support system for the multicriteria analysis of existing stock 196 (2017) 682–689.

55. K. McGlinn, B. Yuce, H. Wicaksono, S. Howell, Y. Rezgui, Usability evaluation of a web-based tool for supporting holistic building energy management, Autom. Constr. 84 (August) (2017) 154–165

56. M. Juszczyk, Implementation of the ANNs ensembles in macro-BIM cost estimates of buildings' floor structural frames, AIP Conf. Proc. 1946 (April) (2018).

57. M. Juszczyk, Cost Estimates of Buildings' Floor Structural Frames with the Use of Support Vector Regression, *IOP Conference Series: Earth and Environmental Science*, vol. 222, no. 1, 2019.

58. T. Su, H. Li, and Y. An, A bim and machine learning integration framework for automated property valuation, J. Build. Eng., vol. 44, 2021.

59. D. Singer, M. Bugler, A. Borrmann. Knowledge based bridge engineering - artificial intelligence meets building information modeling. 2016.