AI-based Resilience for IoT-enabled Business Processes

P Radha Krishna
Dept. of CSE, NIT Warangal, India
prkrishna@nitw.ac.in

Sairam U
Dept. of IT, CBIT, Hyderabad, India
usairam_it@cbit.ac.in

Kamalakar Karlapalem
DSA Centre, IIIT-Hyderabad, India
kamal@iiit.ac.in

Introduction

Business Processes that integrate IoTs in their decision-making need resilience. These processes must be resilient to missing data, incorrect data, uncertain data, etc. For example, when there is missing data, the IoTs cannot determine the situation for decision-making. Thus, Business Process Management (BPM) must be augmented with Artificial Intelligence (AI) techniques to enhance resilience to data issues when dealing with IoTs. IoTs provide the required and relevant data/information from appropriate sensors to execute business processes. We present the problem in the context of missing data, but a closer integration of AI and BPM community to address the problem and enhance the resilience of the IoT-enabled Business Processes is required (see the dotted portion in Figure 1). Execution of Business Processes may fail if the required data is incorrect or unavailable. A closer integration of Business Processes, IoTs, and AI helps in making the business processes more resilient regarding the data and business process failures and results in improved decision-making.

In data analytics, uncertainty is common due to the failure of communication channels or IoT devices. Data flows, originating from a sensor, have a greater impact on the final decision when there are missing values and thereby the Quality-of-Service degrades or undermines the actions to be performed through the IoT-enabled processes (Hossain, et. al., 2018). This necessitates that missing data is either removed or imputed. Most of the time, lost data is imputed using data imputation techniques (Ivan, et. al., 2021). However, the lost data may have important information for activity recognition. On the other hand, in a missing data environment, activity class (types) cannot be estimated when there is no sufficient data available. To improve accuracy during decision-making, it is important to overcome the effects of missing data. Most of the state-of-the-art methods (Straczkiewicz, et. al., 2021) recognize the activities after evaluating the activity data through various machine learning algorithms. The training dataset comprising sensor data

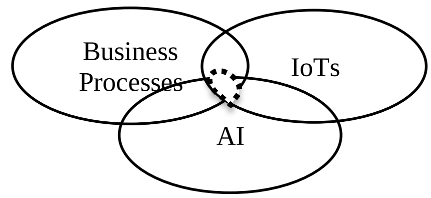

Figure 1: Integration of BP, IoTs and AI

is used to estimate the activity labels. In our work (Sairam et al, 2023), we develop a methodology to handle missing values without imputation so that the overall IoT process does not affect due to insufficient data. We generate subspace decision tree classifiers, where each decision tree uses non-overlapping subsets of feature sets. We build an ensemble classifier over these subspace classifiers to get better performance. We consider different levels of missing data in our ensemble classifier in order to show that it is resilient to missing data and can also give better performance even in cases of incomplete data (that is, data with missing values). We deploy a classical AI-based decision tree to handle the missing data elegantly.

Decision-Making for HBR

Figure 2 shows the part of Human Behaviour Recognition (HBR) business process. The sensor data undergo (pre-)processing steps such as cleaning, transformation, and segmentation and then fuse the sensor data to determine human behaviour. The data captured in a simulation environment is used to train a classifier for the prediction of human behaviour from the data captured from various sensors. Activity classification is the process of associating extracted features with particular activity classes using supervised learning algorithms which are trained to recognize the patterns in the training dataset. This trained model is tested with new observations. The performance of the classification model suffers when data from one or more sensors is missing at certain time periods.

The existing classification approaches handle the missing data by imputation. However, based on the nature of sen-

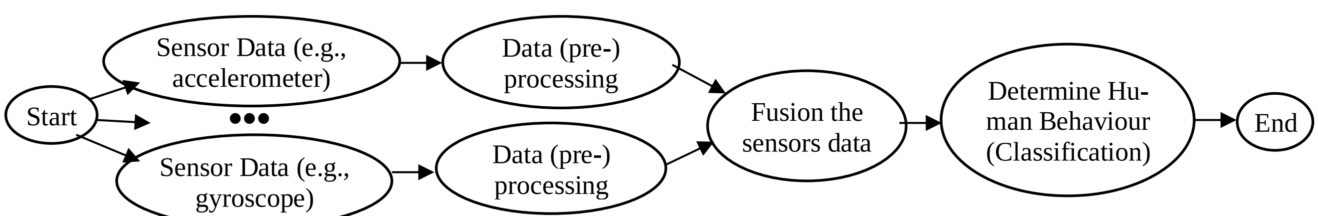

Figure 2: Business process of determining human behaviour from sensors

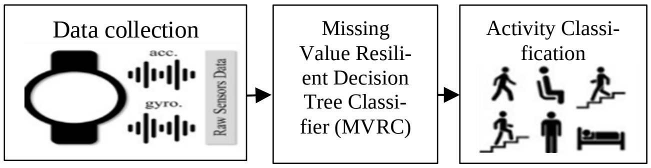

Figure 3: Activity Classification using Sensor data

sor data, imputation may result in unexpected human behavior (Ivan, et. al. 2021). In the next section, we present a specific *missing value resilient classifier* (MVRC) that improves the classification performance even when some data flows from a sensor(s) are missing (Sairam et. al., 2023).

MVRC Approach

Figure 3 shows the high-level overview of the activity classification using sensor data. Our MVRC approach recognizes the activity even when sensor data is lost. Figure 4 shows an overview of the MVRC classifier. Training the MVRC classifier involves building a decision tree on the given dataset, and then adding subspace classifiers built on the features that were not present in previously generated decision trees. We create a supplemented dataset by adding these subspace classifiers' predicted values to the original dataset. The trained subspace decision tree classifiers can serve as an ensemble classifier and handles missing data in the test dataset. The steps involved in constructing the MVRC classifier are: (i) Using training data, construct an iterative convergent subspace decision tree (DT) classifier, (ii) Iterate this process until the attribute list of DT remains unchanged, (iii) Once the iterative non-overlapping subspace classifiers are generated, the outputs of all of them are combined with original features for building a supplemented dataset, and (iv) Build an ensemble classifier trained on the supplemented dataset.

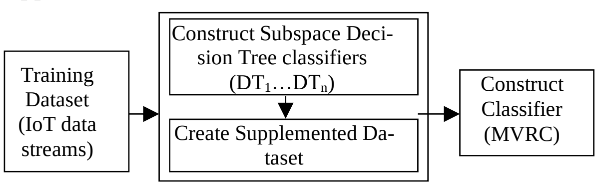

Figure 4: Overview of MVRC classifier

The Motion dataset (https://www.kaggle.com/datasets/outofskills/driving-behavior) consists of data related to the sensors of the Gyroscope and Accelerometer which are used to find the driving behaviour of drivers to avoid road accidents (Wawage and Deshpande, 2022). Aggressive driving includes speeding, sudden left or right turns, and sudden breaks. All of these can be recorded using the sensors in a smartphone. Using the smartphone-embedded accelerometer and gyroscope, the data is captured with 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The sensor signals were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window). We trained the MVRC classifier on a complete dataset and tested with various percentages of missing values in the test dataset.

We considered four example objects from 50% missing values of the dataset (see Table 1). Here, the missing values are generated in attributes 1, 3, 4, and 5 of the first object; attributes 3, and 4 in the second object; attributes 4, and 6 in the third object; and attributes 0, 2, and 5 for the fourth object. The trained ensemble subspace classifiers are applied to these sets of attributes. For the four objects, based on the missing values, our approach chooses the three trained subspace classifiers; and classifies the first, second, third, and fourth objects as Normal, Aggressive, Slow, and Aggressive respectively, which are matching with the ground truth. In general, when missing values are present, these missing values interrupt in proper classification of the decision trees (incorrect decision tree induction and thereby misclassifying). Our MVRC classifier overcomes misclassification. By applying our approach, we can address the presence of missing data in the activity/behavior recognition without data imputation and classify the activities effectively.

Table 1: 50% missing data from 4 example objects

AccX (0)	AccY (1)	AccZ (2)	GyroX (3)	GyroY (4)	GyroZ (5)	Timestamp(6)	Class
0.758194		0.457263				818922.0	Normal
0.667560	-0.038610	0.231416			0.225257	818923.0	Aggressive
2.724449	-7.584121	2.390926	0.023824		-0.038026		Slow
	-0.68167		0.042913	.00366		818923	Aggressive

Discussion and Open Issues

Business Processes dealing with IoT applications face challenges of missing data. AI-based techniques can help alleviate this problem by incorporating the resilience towards missing data by using multiple subspace solutions and developing an ensemble solution over it. We state that without classical AI-driven decision trees, it would have been difficult to get a comprehensive solution to this problem. Our work would help the researchers to develop AI-driven resilient business process solutions for complex IoT business processes with underlying uncertainty in data availability and completeness. There is a need of integrating the three technologies namely BPM, IoTs, and AI for addressing a new class of problems and developing innovative solutions. Some of the open issues and new research directions are:

- Determining cases of failures in IoT business processes by using causality AI solutions detection.
- Enhancing resilience of IoT business processes by incorporating AI techniques.
- Continuous monitoring of IoT business processes to predict and prevent potential failure by using predictive analytics.
- Optimized (or re-specified) data flows, control flows, and event flows for business process automation in IoT applications.
- AI-driven state transfer for replacement sensors to optimize storage, energy, and execution for efficient IoT business processes.

References

Ferrari, A.; Micucci, D.; Mobilio, M.; and Paolo, N. 2021. Trends in human activity recognition using smartphones. *J. Reliable Intell. Environ.* 7, pp.189–213.

Hossain, T.; Goto, H.; and Inoue, S. 2018. Improving Activity Recognition for Missing Data. In Proceedings of the EAI International Conference on Mobile Computing, Applications and Services (MobiCASE) Student Workshop, Osaka, Japan.

Ivan, M. P.; Faisal, H.; Gonçalo, M.; and Nuno, M. G. 2021. Comparison of machine learning techniques for the identification of human activities from inertial sensors available in a mobile device after the application of data imputation techniques, *Computers in Biology and Medicine*. 135. ISSN 0010-4825.

Sairam, U.; Krishna, P. R.; and Kamalakar K. 2023. Missing Data Resilient Ensemble Subspace Decision Tree Classifier, In Proc. of 6th Joint CODS-COMAD 2023, Mumbai, India.

Straczkiewicz, M.; James, P.; and Onnela, JP. 2021. A systematic review of smartphone-based human activity recognition methods for health research. *npj Digit. Med.* 4(148). https://doi.org/10.1038/s41746-021-00514-4

Wawage, P.; and Deshpande, Y. 2022. Smartphone Sensor Dataset for Driver Behavior Analysis. *Data Brief.* 41:107992. doi: 10.1016/j.dib.2022.107992.