Predictive Data Analysis and Machine Learning for Telematics hub based on sensory data

Conference Paper · October 2020

DOI: 10.33107/ubt-ic.2020.214

2 authors:

Dashmir Istrefi
University for Business and Technology

Eftim Zdravevski
Ss. Cyril and Methodius University in Skopje

Predictive Data Analysis and Machine Learning for Telematics Hub based on sensory data

Dashmir Istrefi[1] and Eftim Zdravevski[2]

[1] Information Systems, College for Business and Technology, 10000 Prishtina, Republic of Kosovo, dashmir.istrefi@ubt-uni.net,
[2] Faculty of Computer Science and Engineering, Sts. Cyril and Methodius University in Skopje, Macedonia, eftim.zdravevski@finki.ukim.mk

Abstract. The Internet of things (IoT) paradigm that enables devices to communicate via the Internet, is being adopted in various areas, including smart homes and smart cities. 'Connected Car' is a new term often associated with cars and other passenger vehicles, capable of internet connectivity and sharing various kinds of data with other systems. This paper evaluates the feasibility and technology readiness for this technology to be adopted in the automotive industry. We evaluate different fleet management systems that are currently considered as a state-of-the-art from a scientific point of view. Likewise, we consider commercial platforms and sensory-enabled solutions that streamline the day-to-day operations of different companies.

Keywords: Smart Car, Smart City, IoT, Telematics Hub, Machine Learning, Predictive Maintenance.

1 Introduction

Internet of Things (IoT) is a fast emerging paradigm that is becoming an omnipresent necessity, including in the automobile industry. The data being transmitted can be about the car's location and speed, different parts or lubricants' status, if the car needs urgent service or not, etc. In this sense, the data collected from the devices and IoT sensors can be processed to analyze the driving habits and provide suggestions and optimizations for improving the vehicle's economy and maintenance intervals. The processing of this relies on the fusion of heterogeneous multi-modal [1] and multi-source data and modeling intangible factors, facilitating predictive maintenance and energy saving.

High transport efficiency is also important in today's traffic as haulage is a low margin business with a high turnover. Profit can easily turn into loss by unexpected changes in external conditions such as fuel prices, economic downturns or vehicle failures [2]. By continuously monitoring transport efficiency companies can increase competitiveness and become more profitable. This can be achieved with the use of different Fleet Management Solutions [3]. Fleet management is an operational and strategical challenge to companies and governmental agencies related to passengers and freight transportation services [4]. At an algorithmic level, it relates to optimization and scheduling optimization problems, such as vehicle routing and scheduling, that are notoriously difficult to solve, especially in dynamic settings.

Nowadays, fleet management is based on cloud solutions [5] that need to be combined with combined with approaches for minimizing the total cost of ownership of the cloud infrastructure [6, 7, 8]. A further step is to optimize the collected Big Data to gain insightful information that can improve the business processes and in turn, generate reliable profit streams [9, 10].

The remainder of this paper is structured as follows. Section 2 reviews different commercial fleet management systems as well as academic state-of-the-art approaches. In section 3 we present the architecture of the proposed system and in section 4 we show the preliminary results of our work. Finally, section 5 concludes the paper.

2 Related work

In [11] the authors work on AI powered Collision Prevention system that will not tolerate fatalities caused by machines, taping challenges that guarantee zero accidents. To set the foundation of our research we prepared a set of IoT devices and sensors to be used, it includes a device that supports SDK (Software Development Kit) scripting to prepare unique scenarios and algorithms, Advanced Driver-Assistance Systems (ADAS) [12] and other peripherals for fuel monitoring, temperature, driver identification (see Fig. 1).

We analyzed several Fleet Management Systems covered in scientific publications, such as [13, 14, 15], commercial platforms, with the research module in this paper acting as platform independent, with examples like Geotab [16] and Gurtam [17] and sensory enabled ADAS solution [12]. These solutions streamline the day-to-day operations of any type of company by offering services like route and driver planning, maintenance and repair planning. The biggest challenges outlined by the whitepaper report of Mobileye [18] include a list of what are the biggest headaches when it comes to Fleet Managers:

- Distracted driving – with data from studies suggesting that 71% of truck crashes and 46% of near crashes involve distraction from direct driving related activities as a contributing factor. Findings from Virginia Tech Transportation Institute [19] show that texting while driving increases 23 times more the likelihood of crashing or near crashing.
- Driver retention – with a whitepaper report from Volvo stating that 82% of companies having issues attracting quality drivers [18].
- Busted budgets – apart from fuel making 21% of average fleets marginal cost [20], driver behavior – a category that includes speeding, harsh braking, turn and rapid acceleration, can decrease gas mileage by up to 40%, simulations running through different studies support it [12, 21].
- Speeding Drivers – Vehicles run at higher speeds are more prone to wear and tear. Whether is a new brake pad, tires and engine transmission, maintenance costs add up when vehicle has to be taken off the road for repairs [18].

Fig. 1. Overview of devices and sensors used for research set-up

- Urban driving – Managing a fleet in urban environment involves navigating drivers around bumper to bumper traffic and other inconveniences.

Vehicle uptime is getting increasingly important as transport solutions become more complex and the transport industry seeks new ways of being competitive. Traditional FMS are gradually extended with new features to improve reliability, such as better maintenance planning. Traditional diagnostics methods are unfeasible as it would require too much engineering resources [22]. This paper aims to establish a method (unsupervised and supervised) for predicting vehicle maintenance based on the sensory data.

For a variety of fleets, including intercity trains, school or city buses, taxis, long-haul or short-haul trucks [23], the challenges include ineffective or unnecessary repairs, non-integrated data sources preventing optimizations of the fleet's maintenance, etc. These challenges increase the fleet's operational costs, and cause costly delays [24]. This entails Real-time Fleet Optimization considering the transit network and its schedule [25].

Furthermore, AI can anticipate an engine problem and report it before the driver notices, such as a need of oil change [26]. Such and predictive targeted repairs are only possible with AI-based IoT solutions for data analytics.

The AI-based predictive algorithms can detect problems with equipment, cargo or passengers and providing timely suggestions [27]. AI can analyze routes and incorporate data to give drivers real-time alternative routes in cases of unplanned events [25]. In addition, new technologies such as video-based onboard safety monitoring systems use IoT and AI to help drivers make better decisions [28].

These systems and approaches, demonstrate that this is an attractive field for research and commercialization. Still, there are challenges that are not sufficiently explored or addressed. Therefore, our research aims to implement a system that will help answer these questions:

- Can emerging vehicle malfunctions be detected with AI/ML based on IoT data in order to make pro-active service repairs to avoid unplanned downtime?
- Does the fleet malfunctions and defects correlate with driving styles?
- Can the driving routes be optimized with advanced clustering and optimization algorithms?

3 System Architecture

The implementation of the proposed system consists of the following steps:

- Set up the ecosystem
- Data collection approach: Gather Empirical data to carry out experiments
- Gathering information and data about functionalities. One example of functionality would be predictive maintenance, vehicle break-down, etc.
- Process data collected from the telematics hub and sensors.
- Other forms of semi-structured data obtained by observations and interviews.

The foundation for the ecosystem is hoyo.ai (see Fig. 2 for illustration) that is embedded in SaaS functionalities of different IoT platforms. For research purposes Flespi is used, as an Innovative backend platform-as-a-service (PaaS) specially designed to work with telematics data. It is generally available, providing unified data communication via REST API and a commercial SaaS application with more than 2.3 million connected telematics units worldwide.

The method is data driven and use of extensive amounts of data, either streamed, on-board data or even historic and aggregated data from off-board databases. The methods rely on telematics hub that communicates with IoT platform. A knowledge base is created so that it can be used to predict upcoming failures on other vehicles that show the same deviation. A classifier is trained to learn patterns in the usage data that precede specific repairs and thus can be used to predict vehicle maintenance. Set-up of SDK script to get sensory data from ADAS that records DATA from all types of warnings received from the system such like Lane Departure Warning, Forward Collision Warning.

4 Results

Historical data streams, as shown in Figures 3 and 4, show driver behavior. They can be used a data set to predict potential vehicle malfunctions as some patterns, such as speeding are quite obvious (see Fig. 4).

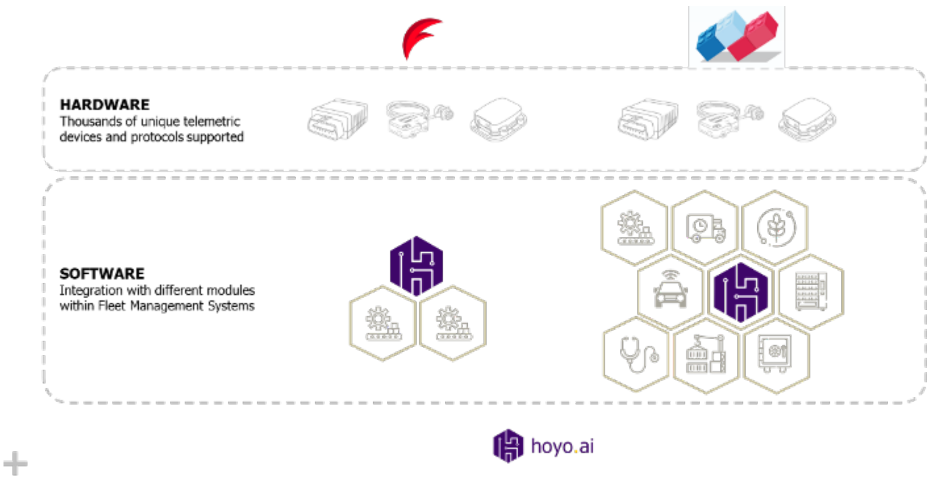

Fig. 2. hoyo.ai architecture

Grouping	Beginning	End	Avg speed	Max speed	Penalties	Violation	Rank	Rating by violations	Duration	Mileage
[redacted]	2019-10-25 07:41:01	2019-10-25 13:01:50	69 km/h	92 km/h	327	-----	2.5	5	1:50:25	43 km
[redacted]	2019-10-25 13:01:50	2019-10-25 18:38:22	62 km/h	86 km/h	20	-----	5.6	1	4:40:35	92 km
[redacted]	2019-10-25 09:03:01	2019-10-25 13:55:15	93 km/h	136 km/h	523		1.9	8	2:02:42	73 km
[redacted]	2019-10-25 08:43:53	2019-10-25 17:52:59	77 km/h	137 km/h	332	-----	2.5	6	5:35:45	129 km
[redacted]	2019-10-25 09:46:10	2019-10-25 16:30:44	60 km/h	73 km/h	35	-----	5.4	2	3:10:41	69 km
[redacted]	2019-10-25 06:59:16	2019-10-25 14:28:09	74 km/h	86 km/h	54	-----	4.9	3	3:44:39	216 km
[redacted]	2019-10-25 08:04:45	2019-10-25 13:25:56	63 km/h	90 km/h	290	-----	2.7	4	1:41:27	49 km
[redacted]	2019-10-25 09:47:45	2019-10-25 16:57:29	71 km/h	102 km/h	536	-----	1.9	9	3:21:57	82 km
[redacted]	2019-10-25 06:47:45	2019-10-25 16:14:18	74 km/h	93 km/h	469	-----	2.1	7	6:35:54	415 km

Fig. 3. Summary historical data streams – taken from implementation on a company in Macedonia

Grouping	Beginning	End	Avg speed	Max speed	Penalties	Violation	Rank	Rating by violations	Duration	Mileage
[redacted]	2020-10-20 07:46:57	2020-10-24 18:17:03	27 km/h	93 km/h	29.446244	-----	5.6	3	18:30:25	492 km
2020-10-20 09:32:11	2020-10-20 09:32:11	2020-10-20 09:32:24	63 km/h	65 km/h	100	Speeding: mild	0	-----	0:00:13	0.23 km
2020-10-20 12:58:22	2020-10-20 12:58:22	2020-10-20 12:58:48	91 km/h	91 km/h	100	Speeding: mild	0	-----	0:00:26	0.65 km
2020-10-20 15:02:17	2020-10-20 15:02:17	2020-10-20 15:03:05	55 km/h	60 km/h	100	Speeding: mild	0	-----	0:00:48	0.73 km
2020-10-20 17:54:08	2020-10-20 17:54:08	2020-10-20 17:54:40	80 km/h	85 km/h	2000	Speeding: medium	0	-----	0:00:32	0.71 km
2020-10-21 13:28:46	2020-10-21 13:28:46	2020-10-21 13:29:32	69 km/h	72 km/h	100	Speeding: mild	0	-----	0:00:46	0.88 km
2020-10-21 13:30:07	2020-10-21 13:30:07	2020-10-21 13:30:41	56 km/h	60 km/h	100	Speeding: mild	0	-----	0:00:34	0.53 km
2020-10-21 15:09:25	2020-10-21 15:09:25	2020-10-21 15:09:49	68 km/h	68 km/h	100	Speeding: mild	0	-----	0:00:24	0.45 km
2020-10-21 15:27:34	2020-10-21 15:27:34	2020-10-21 15:28:00	58 km/h	63 km/h	100	Speeding: mild	0	-----	0:00:26	0.42 km
2020-10-21 17:13:21	2020-10-21 17:13:21	2020-10-21 17:13:58	64 km/h	66 km/h	100	Speeding: mild	0	-----	0:00:37	0.66 km
2020-10-21 17:14:22	2020-10-21 17:14:22	2020-10-21 17:15:07	61 km/h	66 km/h	100	Speeding: mild	0	-----	0:00:45	0.76 km
2020-10-21 17:17:46	2020-10-21 17:17:46	2020-10-21 17:19:11	59 km/h	63 km/h	100	Speeding: mild	0	-----	0:01:25	1.39 km
2020-10-22 14:52:17	2020-10-22 14:52:17	2020-10-22 14:52:44	58 km/h	61 km/h	100	Speeding: mild	0	-----	0:00:27	0.43 km
2020-10-22 15:37:06	2020-10-22 15:37:06	2020-10-22 15:37:17	60 km/h	60 km/h	100	Speeding: mild	0	-----	0:00:12	0.20 km
2020-10-22 16:15:51	2020-10-22 16:15:51	2020-10-22 16:16:04	59 km/h	63 km/h	100	Speeding: mild	0	-----	0:00:13	0.21 km
2020-10-22 17:09:11	2020-10-22 17:09:11	2020-10-22 17:09:24	60 km/h	63 km/h	100	Speeding: mild	0	-----	0:00:13	0.22 km
2020-10-22 17:10:07	2020-10-22 17:10:07	2020-10-22 17:10:33	60 km/h	62 km/h	100	Speeding: mild	0	-----	0:00:26	0.43 km
2020-10-23 08:53:48	2020-10-23 08:53:48	2020-10-23 08:54:16	86 km/h	93 km/h	100	Speeding: mild	0	-----	0:00:28	0.67 km
2020-10-23 08:54:21	2020-10-23 08:54:21	2020-10-23 08:56:09	70 km/h	83 km/h	2000	Speeding: medium	0	-----	0:01:48	2.09 km
2020-10-23 09:50:44	2020-10-23 09:50:44	2020-10-23 09:51:55	60 km/h	63 km/h	100	Speeding: mild	0	-----	0:01:11	1.18 km
2020-10-23 15:11:37	2020-10-23 15:11:37	2020-10-23 15:11:49	67 km/h	66 km/h	100	Speeding: mild	0	-----	0:00:12	0.22 km
2020-10-23 16:19:24	2020-10-23 16:19:24	2020-10-23 16:19:56	61 km/h	77 km/h	2000	Speeding: medium	0	-----	0:00:32	0.56 km
2020-10-24 17:08:17	2020-10-24 17:08:17	2020-10-24 17:08:29	68 km/h	71 km/h	2000	Speeding: medium	0	-----	0:00:12	0.23 km
2020-10-24 17:24:27	2020-10-24 17:24:27	2020-10-24 17:24:34	38 km/h	65 km/h	300	Harsh driving	0	-----	0:00:07	0.07 km
2020-10-24 17:24:31	2020-10-24 17:24:31	2020-10-24 17:24:34	20 km/h	14 km/h	2000	Brake: extreme	0	-----	0:00:03	0.02 km
2020-10-24 17:25:28	2020-10-24 17:25:28	2020-10-24 17:25:41	60 km/h	64 km/h	100	Speeding: mild	0	-----	0:00:13	0.22 km
2020-10-24 17:27:32	2020-10-24 17:27:32	2020-10-24 17:28:05	71 km/h	72 km/h	100	Speeding: mild	0	-----	0:00:33	0.65 km
2020-10-24 17:45:28	2020-10-24 17:45:28	2020-10-24 17:46:06	71 km/h	72 km/h	100	Speeding: mild	0	-----	0:00:38	0.75 km
2020-10-24 17:58:49	2020-10-24 17:58:49	2020-10-24 17:59:11	68 km/h	74 km/h	2000	Speeding: medium	0	-----	0:00:22	0.42 km
2020-10-24 18:07:10	2020-10-24 18:07:10	2020-10-24 18:07:50	48 km/h	51 km/h	100	Speeding: mild	0	-----	0:00:40	0.53 km
[redacted]	2020-10-19 11:31:47	2020-10-24 18:19:51	24 km/h	133 km/h	108.310708	-----	3.9	4	1 days 0:38:51	598 km
[redacted]	2020-10-21 10:08:17	2020-10-24 13:53:43	47 km/h	57 km/h	17.936611	-----	5.9	1	2:21:39	112 km
[redacted]	2020-10-20 13:12:39	2020-10-24 09:21:56	30 km/h	139 km/h	19.308047	-----	5.9	2	16:29:17	492 km
Total	2020-10-19 11:31:47	2020-10-24 18:19:51	27 km/h	139 km/h	53.593727	-----	4.9	-----	2 days 14:00:12	1694 km

Fig. 4. Detailed historical data streams – taken from implementation on a company in Macedonia

With the use of machine learning on other sensory [29], we believe that learnings from these streams are quite feasible as well. We intend to use a self-organized approach to detect deviations. It will rely on a database built of deviations known to later cause unplanned repairs by matching the disappearance of deviations to a documented repair. The dataset can be also used to predict failures as already known deviations reappear in different vehicles. The method could enable a life-long learning as the database gradually evolves over time and new deviations are learned. With that, the predictive maintenance could evolve over time and react to unforeseen phenomena or failures. The method could be cost effective as it requires little manual resources and easily expandable to new vehicle components.

The expected outcomes of this research are:

- Find the appropriate algorithms that could identify the causes of vehicle malfunctions and faults based on the historic pattern of sensory measurements.
- Failures related to cooling fan, or heat load of the engine.
- Estimate driving styles based on wheel speed sensors, bakes, turn, acceleration and considering road conditions (rural, urban).
- Evaluate the approach in real-life setting of 1000 telematic hubs divided into different categories (light vehicles, trucks, buses, heavy machinery).
- Integrate the hoyo.ai modules with established IoT platforms, thus offering it to 3+ million units worldwide.
- Perform cost/benefit analysis of using the predictive maintenance algorithms.

5 Conclusion

This paper reviewed current state-of-the-art for predictive maintenance in the automotive industry. Likewise, it analyzed the offerings of current commercial platforms. Based on the identified shortcomings of these approaches, it proposed a novel platform, hoyo.ai, to integrate sensors and collect data. Based on the analysis preliminary collected data based on a deployment in a company in Macedonia, we can conclude that the patterns in the data are evident. In the future, we intend to evaluate different classical machine learning and deep learning approaches for automatically analyzing these data sets, which are growing by the day as data is being collected with hoyo's current and new customers.

References

[1] Kalajdjieski, J., Zdravevski, E., Corizzo, R., Lameski, P., Kalajdziski, S., Pires, I.M., Garcia, N.M., Trajkovik, V.: Air pollution prediction with multi-modal data and deep neural networks. Remote Sensing **12**(24) (2020) 4142

[2] Prytz, R.: Machine learning methods for vehicle predictive maintenance using off-board and on-board data. PhD thesis, Halmstad University Press (2014)

[3] Moradi Afrapoli, A., Askari-Nasab, H.: Mining fleet management systems: a review of models and algorithms. International Journal of Mining, Reclamation and Environment **33**(1) (2019) 42–60

[4] Bielli, M., Bielli, A., Rossi, R.: Trends in models and algorithms for fleet management. Procedia-Social and Behavioral Sciences **20** (2011) 4–18

[5] Xu, G., Li, M., Luo, L., Chen, C.H., Huang, G.Q.: Cloud-based fleet management for prefabrication transportation. Enterprise Information Systems **13**(1) (2019) 87–106

[6] Grzegorowski, M., Zdravevski, E., Janusz, A., Lameski, P., Apanowicz, C., Slezak, D.: Cost optimization for big data workloads based on dynamic scheduling and cluster-size tuning. Big Data Research **25** (2021) 100203

[7] Zdravevski, E., Lameski, P., Dimitrievski, A., Grzegorowski, M., Apanowicz, C.: Cluster-size optimization within a cloud-based etl framework for big data. In: 2019 IEEE International Conference on Big Data (Big Data). (2019) 3754–3763

[8] Zdravevski, E., Lameski, P., Kulakov, A.: Row Key Designs of NoSQL Database Tables and Their Impact on Write Performance. In: Proceedings - 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, PDP 2016, IEEE (feb 2016) 10–17

[9] Zdravevski, E., Lameski, P., Kulakov, A., Filiposka, S., Trajanov, D., Jakimovski, B.: Parallel computation of information gain using hadoop and mapreduce. In M. Ganzha, L. Maciaszek, M.P., ed.: Proceedings of the 2015 Federated Conference on Computer Science and Information Systems. Volume 5 of Annals of Computer Science and Information Systems., IEEE (2015) 181–192

[10] Zdravevski, E., Lameski, P., Apanowicz, C., Slezak, D.: From big data to business analytics: The case study of churn prediction. Applied Soft Computing **90** (2020) 106164

[11] Shalev-Shwartz, S., Shammah, S., Shashua, A.: On a formal model of safe and scalable self-driving cars. mobileye. arXiv preprint arXiv:1708.06374 (2017)

[12] Zeman, J.: Adas and fuel economy: Competing objectives? `https://www.gtisoft.com/adas-and-fuel-economy-competing-objectives/` (2020-06-08) Accessed on 2020-09-01.

[13] Reimer, M.: Service relationship management–driving uptime in commercial vehicle maintenance and repair. DECISIV, White paper (2013)

[14] Fornaia, A., Napoli, C., Pappalardo, G., Tramontana, E.: Enhancing city transportation services using cloud support. In: International Conference on Information and Software Technologies, Springer (2016) 695–708

[15] Steiner, J.C., Armstrong, C., Kress, T., Walli, T., Gallagher, R.J., Ngo, J., Silva, A.: Commercial vehicle global positioning system based telematics data characteristics and limitations. Technical report, SAE Technical Paper (2017)

[16] Sherzad, O.: Geotab. collision reconstruction with telematics. https://www.geotab.com/white-paper/collision-reconstruction-with-telematics/ (2020-06-08) Accessed on 2020-09-01.

[17] : Preventing road accidents with wialon in russia. use case. https://gurtam.com/en/case-studies/preventing-road-accidents Accessed on 2020-09-01.

[18] : The 5 biggest headaches fleet managers face, and how to deal with them. https://static.mobileye.com/website/au/fleets/files/White%20paper%205%20Biggest%20Headaches%20Fleet%20Managers%20Face.pdf Accessed on 2020-09-01.

[19] : Virginia tech transportation institute (vtti): Impacts transportation policy. https://featured.vtti.vt.edu/?p=197 Accessed on 2020-09-01.

[20] Hooper, A., Murray, D.: An analysis of the operational costs of trucking: 2018 update. (2018)

[21] Istrefi, D.: Innovation in transport and logistics through the use of technology: A case study. In: ICMBE Conferences. (2019)

[22] Magargle, R., Johnson, L., Mandloi, P., Davoudabadi, P., Kesarkar, O., Krishnaswamy, S., Batteh, J., Pitchaikani, A.: A simulation-based digital twin for model-driven health monitoring and predictive maintenance of an automotive braking system. In: Proceedings of the 12th International Modelica Conference, Prague, Czech Republic, May 15-17, 2017. Number 132, Linköping University Electronic Press (2017) 35–46

[23] Raa, B.: Fleet optimization for cyclic inventory routing problems. International Journal of Production Economics **160** (2015) 172–181

[24] Ströhle, P., Flath, C.M., Gärttner, J.: Leveraging customer flexibility for car-sharing fleet optimization. Transportation Science **53**(1) (2019) 42–61

[25] Guihaire, V., Hao, J.K.: Transit network design and scheduling: A global review. Transportation Research Part A: Policy and Practice **42**(10) (2008) 1251–1273

[26] Jun, H.B., Kiritsis, D., Gambera, M., Xirouchakis, P.: Predictive algorithm to determine the suitable time to change automotive engine oil. Computers & Industrial Engineering **51**(4) (2006) 671–683

[27] Yigitcanlar, T., Desouza, K.C., Butler, L., Roozkhosh, F.: Contributions and risks of artificial intelligence (ai) in building smarter cities: Insights from a systematic review of the literature. Energies **13**(6) (2020) 1473

[28] Camden, M.C., Medina-Flintsch, A., Hickman, J.S., Miller, A.M., Hanowski, R.J.: Leveraging large-truck technology and engineering to realize safety gains: Video-based onboard safety monitoring systems. (2017)

[29] Zdravevski, E., Lameski, P., Trajkovik, V., Kulakov, A., Chorbev, I., Goleva, R., Pombo, N., Garcia, N.: Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering. IEEE Access **5** (2017) 5262–5280