Verification and Validation of Systems in Which AI is a Key Element

Verification and Validation of Systems in Which AI is a Key Element

---

Lead Author: *Laura Pullum*

Many systems are being considered in which artificial intelligence (AI) will be a key element. Failure of an AI element can lead to system failure (Dreossi et al 2017), hence the need for AI verification and validation (V&V). The element(s) containing AI capabilities is treated as a subsystem and V&V is conducted on that subsystem and its interfaces with other elements of the system under study, just as V&V would be conducted on other subsystems. That is, the high-level definitions of V&V do not change for systems containing one or more AI elements.

However, AI V&V challenges require approaches and solutions beyond those for conventional or traditional (those without AI elements) systems. This article provides an overview of how machine learning components/subsystems "fit" in the systems engineering framework, identifies characteristics of AI subsystems that create challenges in their V&V, illuminates those challenges, and provides some potential solutions while noting open or continuing areas of research in the V&V of AI subsystems.

☐

Contents

Overview of V&V for AI-based Systems

Conventional systems are engineered via 3 overarching phases, namely, requirements, design and V&V. These phases are applied to each subsystem and to the system under study. As shown in Figure 1, this is the case even if the subsystem is based on AI techniques.

Error creating thumbnail: File missing

Figure 1. Systems Engineering Phases for Systems Containing Machine Learning and Conventional Subsystems. (SEBoK Original, modeled after (Kuwajima et al. 2020))

AI-based systems follow a different lifecycle than do traditional systems. As shown in the general machine learning life cycle illustrated in Figure 2, V&V activities occur throughout the life cycle. In addition to requirements allocated to the AI subsystem (as is the case for conventional subsystems), there also may be requirements for data that flow up to the system from the AI subsystem.

Error creating thumbnail: File missing

Figure 2. General AI Life Cycle/Workflow. (SEBoK Original)

Characteristics of AI Leading to V&V Challenges

Though some aspects of V&V for conventional systems can be used without modification, there are important characteristics of AI subsystems that lead to challenges in their verification and validation. In a survey of engineers, Ishikawa and Yoshioka (2019) identify attributes of machine learning that make the engineering of same difficult. According to the engineers surveyed, the top attributes with a summary of the engineers' comments are:

- *Lack of an oracle*: It is difficult or impossible to clearly define the correctness criteria for system outputs or the right outputs for each individual input.
- *Imperfection*: It is intrinsically impossible to for an AI system to be 100% accurate.
- *Uncertain behavior for untested data*: There is high uncertainty about how the system will behave in response to untested input data, as evidenced by radical changes in behavior given slight changes in input (e.g., adversarial examples).
- *High dependency of behavior on training data*: System behavior is highly dependent on the training data.

These attributes are characteristic of AI itself and can be generalized as follows:

- Erosion of determinism
- Unpredictability and unexplainability of individual outputs (Sculley et al., 2014)
- Unanticipated, emergent behavior, and unintended consequences of algorithms
- Complex decision making of the algorithms
- Difficulty of maintaining consistency and weakness against slight changes in inputs (Goodfellow et al., 2015)

V&V Challenges of AI Systems

Requirements

Challenges with respect to AI requirements and AI requirements engineering are extensive and due in part to the practice by some to treat the AI element as a "black box" (Gunning 2016). Formal specification has been attempted and has shown to be difficult for those hard-to-formalize tasks and requires decisions on the use of quantitative or Boolean specifications and the use of data and formal requirements. The challenge here is to design effective methods to specify both desired and undesired properties of systems that use AI- or ML-based components (Seshia 2020).

A taxonomy of AI requirements engineering challenges, outlined by Belani and colleagues (2019), is shown in Table 1.

Table 1: Requirements engineering for AI (RE4AI) taxonomy, mapping challenges to AI-related entities and requirements engineering activities (after (Belani et al., 2019))

RE4AI	AI Related Entities		
RE Activities	**Data**	**Model**	**System**
Elicitation	- Availability of large datasets - Requirements analyst upgrade	- Lack of domain knowledge - Undeclared consumers	- How to define problem /scope - Regulation (e.g., ethics) not clear

Analysis	- Imbalanced datasets, silos - Role: data scientist needed	- No trivial workflows - Automation tools needed	- No integration of end results - Role: business analyst upgrade
Specification	- Data labelling is costly, needed - Role: data engineer needed	- No end-to-end pipeline support - Minimum viable model useful	- Avoid design anti- patterns - Cognitive / system architect needed
Validation	- Training data critical analysis - Data dependencies	- Entanglement, CACE problem - High scalability issues for ML	- Debugging, interpretability - Hidden feedback loops
Management	- Experiment management - No GORE-like method polished	- Difficult to log and reproduce - DevOps role for AI needed	- IT resource limitations, costs - Measuring performance
Documentation	- Data & model visualization - Role: research scientist useful	- Datasets and model versions - Education and training of staff	- Feedback from end-users - Development method
All of the Above	- Data privacy and data safety - Data dependencies		

CACE: change anything, change everything

GORE: goal-oriented requirements engineering

Data

Data is the life-blood of AI capabilities given that it is used to train and evaluate AI models and produce their capabilities. Data quality attributes of importance to AI include accuracy, currency and timeliness, correctness, consistency, in addition to usability, security and privacy, accessibility, accountability, scalability, lack of bias and others. As noted above, the correctness of unsupervised methods is embedded in the training data and the environment.

There is a question of coverage of the operational space by the training data. If the data does not adequately cover the operational space, the behavior of the AI component is questionable. However, there are no strong guarantees on when a data set it ‘large enough’. In addition, ‘large’ is not sufficient. The data must sufficiently cover the operational space.

Another challenge with data is that of adversarial inputs. Szegedy et al. (2014) discovered that several ML models are vulnerable to adversarial examples. This has been shown many times on image classification software, however, adversarial attacks can be made against other AI tasks (e.g., natural language processing) and against techniques other than neural networks (typically used in image classification) such as reinforcement learning (e.g., reward hacking) models.

Model

Numerous V&V challenges arise in the model space, some of which are provided below.

- *Modeling the environment*: Unknown variables, determining the correct fidelity to model, modeling human behavior. The challenge problem is providing a systematic method of environment modeling that allows one to provide provable guarantees on the system’s behavior even when there is considerable uncertainty about the environment. (Seshia 2020)
- *Modeling learning systems*: Very high dimensional input space, very high dimensional parameter or state space, online adaptation/evolution, modeling context (Seshia 2020).
- *Design and verification of models and data*: data generation, quantitative verification, compositional reasoning, and compositional specification (Seshia 2020). The challenge is to develop techniques for compositional reasoning that do not rely on having complete compositional specifications (Seshia 2017).
- *Optimization strategy must balance between over- and under-specification*. One approach, instead of using distance (between predicted and actual results) measures, uses the cost of an erroneous result (e.g., an incorrect classification) as a criterion (Faria, 2018) (Varshney, 2017).
- *Online learning*: requires monitoring; need to ensure its exploration does not result in unsafe states.
- *Formal methods*: intractable state space explosion

from complexity of the software and the system’s interaction with its environment, an issue with formal specifications.

- *Bias* in algorithms from underrepresented or incomplete training data OR reliance on flawed information that reflects historical inequities. A biased algorithm may lead to decisions with collective disparate impact. Trade-off between fairness and accuracy in the mitigation of an algorithm’s bias.
- *Test coverage*: effective metrics for test coverage of AI components is an active area of research with several candidate metrics, but currently no clear best practice.

Properties

Assurance of several AI system properties is necessary to enable trust in the system, e.g., the system’s trustworthiness. This is a separate though necessary aspect of system dependability for AI systems. Some important properties are listed below and though extensive, are not comprehensive.

- *Accountability*: refers to the need of an AI system to be answerable for its decisions, actions and performance to users and others with whom the AI system interacts
- *Controllability*: refers to the ability of a human or other external agent to intervene in the AI system’s functioning
- *Explainability*: refers to the property of an AI system to express important factors influencing the AI system results or to provide details/reasons behind its functioning so that humans can understand
- *Interpretability*: refers to the degree to which a human can understand the cause of a decision (Miller 2017)
- *Reliability*: refers to the property of consistent intended behavior and results
- *Resilience*: refers to the ability of a system to recover operations quickly following an incident
- *Robustness*: refers to the ability of a system to maintain its level of performance when errors occur during execution and to maintain that level of performance given erroneous inputs and parameters

- *Safety*: refers to the freedom from unacceptable risk
- *Transparency*: refers to the need to describe, inspect and reproduce the mechanisms through which AI systems make decisions, communicating this to relevant stakeholders.

V&V Approaches and Standards

V&V Approaches

Prior to the proliferation of deep learning, research on V&V of neural networks touched on adaptation of available standards, such as the then-current IEEE Std 1012 (Software Verification and Validation) processes (Pullum et al. 2007), areas need to be augmented to enable V&V (Taylor 2006), and examples of V&V for high-assurance systems with neural networks (Schumann et al., 2010). While these books provide techniques and lessons learned, many of which remain relevant, additional challenges due to deep learning remain unsolved.

One of the challenges is data validation. It is vital that the data upon which AI depends undergo V&V. Data quality attributes that are important for AI systems include accuracy, currency and timeliness, correctness, consistency, usability, security and privacy, accessibility, accountability, scalability, lack of bias, and coverage of the state space. Data validation steps can include file validation, import validation, domain validation, transformation validation, aggregation rule and business validation (Gao et al. 2011).

There are several approaches to V&V of AI components, including formal methods (e.g., formal proofs, model checking, probabilistic verification), software testing, simulation-based testing and experiments. Some specific approaches are:

- Metamorphic testing to test ML algorithms, addressing the oracle problem (Xie et al., 2011)
- A ML test score consisting of tests for features and data, model development and ML infrastructure, and monitoring tests for ML (Breck et al., 2016)
- Checking for inconsistency with desired behavior and systematically searching for worst-case outcomes when testing consistency with specifications.
- Corroborative verification (Webster et al., 2020), in

which several verification methods, working at different levels of abstraction and applied to the same AI component, may prove useful to verification of AI components of systems.

- Testing against strong adversarial attacks (Useato, 2018); researchers have found that models may show robustness to weak adversarial attacks and show little to no accuracy to strong attacks (Athalye et al., 2018, Uesato et al., 2018, Carlini and Wagner, 2017).
- Use of formal verification to prove that models are consistent with specifications, e.g., (Huang et al., 2017).
- Assurance cases combining the results of V&V and other activities as evidence to support claims on the assurance of systems with AI components (Kelly and Weaver, 2004; Picardi et al. 2020).

Standards

Standards development organizations (SDO) are earnestly working to develop standards in AI, including the safety and trustworthiness of AI systems. Below are just a few of the SDOs and their AI standardization efforts.

ISO is the first international SDO to set up an expert group to carry out standardization activities for AI. Subcommittee (SC) 42 is part of the joint technical committee ISO/IEC JTC 1. SC 42 has a working group on foundational standards to provide a framework and a common vocabulary, and several other working groups on computational approaches to and characteristics of AI systems, trustworthiness, use cases, applications, and big data. (https://www.iso.org/committee/6794475.html)

The IEEE P7000 series of projects are part of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, launched in 2016. IEEE P7009, "Fail-Safe Design of Autonomous and Semi-Autonomous Systems" is one of 13 standards in the series. (https://standards.ieee.org/project/7009.html)

Underwriters Laboratory has been involved in technology safety for 125 years and has released ANSI/UL 4600 "Standard for Safety for the Evaluation of Autonomous Products". (https://ul.org/UL4600)

The SAE G-34, Artificial Intelligence in Aviation, Committee is responsible for creating and maintaining

SAE Technical Reports, including standards, on the implementation and certification aspects related to AI technologies inclusive of any on or off-board system for the safe operation of aerospace systems and aerospace vehicles. (https://www.sae.org/works/committeeHome.do?comtID=TEAG34)

References

Works Cited

Belani, Hrvoje, Marin Vuković, and Željka Car. Requirements Engineering Challenges in Building AI-Based Complex Systems. 2019. IEEE 27th International Requirements Engineering Conference Workshops (REW).

Breck, Eric, Shanqing Cai, Eric Nielsen, Michael Salib and D. Sculley. What's your ML Test Score? A Rubric for ML Production Systems. 2016. 30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona Spain.

Daume III, Hal, and Daniel Marcu. Domain adaptation for statistical classifiers. *Journal of Artificial Intelligence Research*, 26:101–126, 2006.

Dreossi, T., A. Donzé, S.A. Seshia. Compositional falsification of cyber-physical systems with machine learning components. In Barrett, C., M. Davies, T. Kahsai (eds.) NFM 2017. LNCS, vol. 10227, pp. 357-372. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57288-8_26

Faria, José M. Machine learning safety: An overview. In *Proceedings of the 26th Safety-Critical Systems Symposium*, York, UK, February 2018.

Farrell, M., Luckcuck, M., Fisher, M. Robotics and Integrated Formal Methods. Necessity Meets Opportunity. In: *Integrated Formal Methods*. pp. 161-171. Springer (2018).

Gao, Jerry, Chunli Xie, and Chuanqi Tao. 2016. Big Data Validation and Quality Assurance – Issues, Challenges and Needs. 2016 IEEE Symposium on Service-Oriented System Engineering (SOSE), Oxford, UK, 2016, pp. 433-441, doi: 10.1109/SOSE.2016.63.

Gleirscher, M., Foster, S., Woodcock, J. New

Opportunities for Integrated Formal Methods. *ACM Computing Surveys* 52(6), 1-36 (2020).

Goodfellow, Ian, J. Shlens, C. Szegedy. Explaining and harnessing adversarial examples. In International Conference on Learning Representations (ICLR), May 2015.

Gunning, D. Explainable Artificial Intelligence (XAI). In IJCAI 2016 Workshop on Deep Learning for Artificial Intelligence (DLAI), July 2016.

Huang, X., M. Kwiatkowska, S. Wang, and M. Wu. Safety Verification of deep neural networks. In. Majumdar, R., and V. Kunčak (eds.) CAV 2017. LNCS, vol. 10426, pp. 3-29. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63387-9_1

Ishikawa, Fuyuki and Nobukazu Yoshioka. How do Engineers Perceive Difficulties in Engineering of Machine-Learning Systems? - Questionnaire Survey. 2019 IEEE/ACM Joint 7th International Workshop on Conducting Empirical Studies in Industry (CESI) and 6th International Workshop on Software Engineering Research and Industrial Practice (SER&IP) (2019)

Jones, Cliff B. Tentative steps toward a development method for interfering programs. *ACM Transactions on Programming Languages and Systems* (TOPLAS), 5(4):596–619, 1983.

Kelly, T., and R. Weaver. The goal structuring notation – a safety argument notation. In Dependable Systems and Networks 2004 Workshop on Assurance Cases, July 2004.

Klein, G., Andronick, J., Fernandez, M., Kuz, I., Murray, T., Heiser, G. Formally verified software in the real world. *Comm. of the ACM* 61(10), 68-77 (2018).

Kuwajima, Hiroshi, Hirotoshi Yasuoka, and Toshihiro Nakae. Engineering problems in machine learning systems. *Machine Learning* (2020) 109:1103–1126. https://doi.org/10.1007/s10994-020-05872-w

Lwakatare, Lucy Ellen, Aiswarya Raj, Ivica Crnkovic, Jan Bosch, and Helena Holmström Olsson. Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions. *Information and Software Technology* 127 (2020) 106368

Luckcuck, M., Farrell, M., Dennis, L.A., Dixon, C., Fisher, M. Formal Specification and Verification of Autonomous Robotic Systems: A Survey. *ACM Computing Surveys* 52(5), 1-41 (2019).

Marijan, Dusica and Arnaud Gotlieb. Software Testing for Machine Learning. The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20) (2020)

Miller, Tim. Explanation in artificial intelligence: Insights from the social sciences. arXiv Preprint arXiv:1706.07269. (2017).

Pei, K., Y. Cao, J Yang, and S. Jana. DeepXplore: automated whitebox testing of deep learning systems. In The 26th Symposium on Operating Systems Principles (SOSP 2017), pp. 1-18, October 2017.

Picardi, Chiara, Paterson, Colin, Hawkins, Richard David et al. (2020) Assurance Argument Patterns and Processes for Machine Learning in Safety-Related Systems. In: *Proceedings of the Workshop on Artificial Intelligence Safety* (SafeAI 2020). CEUR Workshop Proceedings, pp. 23-30.

Pullum, Laura L., Brian Taylor, and Marjorie Darrah, *Guidance for the Verification and Validation of Neural Networks*, IEEE Computer Society Press (Wiley), 2007.

Rozier, K.Y. Specification: The Biggest Bottleneck in Formal Methods and Autonomy. In: *Verified Software. Theories, Tools, and Experiments*. pp. 8-26. Springer (2016).

Schumann, Johan, Pramod Gupta and Yan Liu. Application of neural networks in High Assurance Systems: A Survey. In *Applications of Neural Networks in High Assurance Systems*, Studies in Computational Intelligence, pp. 1-19. Springer, Berlin, Heidelberg, 2010.

Sculley, D., Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-François Crespo, and Dan Dennison. Machine Learning: the high interest credit card of technical debt. In NIPS 2014 Workshop on Software Engineering for Machine Learning (SE4ML), December 2014.

Seshia, Sanjit A. Compositional verification without compositional specification for learning-based systems. Technical Report UCB/EECS-2017-164, EECS Department, University of California, Berkeley, Nov 2017.

Seshia, Sanjit A., Dorsa Sadigh, and S. Shankar Sastry. Towards Verified Artificial Intelligence. arXiv:1606.08514v4 [cs.AI] 23 Jul 2020.

Szegedy, Christian, Zaremba, Wojciech, Sutskever, Ilya, Bruna, Joan, Erhan, Dumitru, Goodfellow, Ian J., and Fergus, Rob. Intriguing properties of neural networks. ICLR, abs/1312.6199, 2014b. URL http://arxiv.org/abs/1312.6199.

Taylor, Brian, ed. *Methods and Procedures for the Verification and Validation of Artificial Neural Networks*, Springer-Verlag, 2005.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge, MA: Harvard University Press.

Tiwari, Ashish, Bruno Dutertre, Dejan Jovanović, Thomas de Candia, Patrick D. Lincoln, John Rushby, Dorsa Sadigh, and Sanjit Seshia. Safety envelope for security. In *Proceedings of the 3rd International Conference on High Confidence Networked Systems* (HiCoNS), pp. 85-94, Berlin, Germany, April 2014. ACM.

Uesato, Jonathan, O'Donoghue, Brendan, van den Oord, Aaron, Kohli, Pushmeet. Adversarial Risk and the Dangers of Evaluating Against Weak Attacks. *Proceedings of the 35th International Conference on Machine Learning*, Stockholm, Sweden, PMLR 80, 2018.

Varshney, Kush R., and Homa Alemzadeh. On the safety of machine learning: Cyber-physical systems, decision sciences, and data products. *Big Data*, 5(3):246–255, 2017.

Webster, M., Wester, D.G., Araiza-Illan, D., Dixon, C., Eder, K., Fisher, M., Pipe, A.G. A corroborative approach to verification and validation of human-robot teams. *J. Robotics Research* 39(1) (2020).

Xie, Xiaoyuan, J.W.K. Ho, C. Murphy, G. Kaiser, B. Xu, and T.Y. Chen. 2011. "Testing and Validating Machine Learning Classifiers by Metamorphic Testing," *Journal of Software Testing*, April 1, 84(4): 544-558, doi:10.1016/j.jss.2010.11.920.

Zhang, J., Li, J. Testing and verification of neural-network-based safety-critical control software: A systematic literature review. *Information and Software Technology* 123, 106296 (2020).

Zhang, J.M., Harman, M., Ma, L., Liu, Y. Machine learning testing: Survey, landscapes and horizons. *IEEE Transactions on Software Engineering*. 2020, doi: 10.1109/TSE.2019.2962027.

Primary References

Belani, Hrvoje, Marin Vuković, and Željka Car. Requirements Engineering Challenges in Building AI-Based Complex Systems. 2019. IEEE 27th International Requirements Engineering Conference Workshops (REW).

Dutta, S., Jha, S., Sankaranarayanan, S., Tiwari, A. 2018. Output range analysis for deep feedforward neural networks. In: NASA Formal Methods. pp. 121-138.

Gopinath, D., G. Katz, C. Pāsāreanu, and C. Barrett. 2018. DeepSafe: A Data-Driven Approach for Assessing Robustness of Neural Networks. In: *ATVA*.

Huang, X., M. Kwiatkowska, S. Wang and M. Wu. 2017. Safety Verification of Deep Neural Networks. Computer Aided Verification.

Jha, S., V. Raman, A. Pinto, T. Sahai, and M. Francis. 2017. On Learning Sparse Boolean Formulae for Explaining AI Decisions, *NASA Formal Methods*.

Katz, G., C. Barrett, D. Dill, K. Julian, M. Kochenderfer. 2017. Reluplex: An Efficient SMT Solver for Verifying Deep Neural Networks, https://arxiv.org/abs/1702.01135.

Leofante, F., N. Narodytska, L. Pulina, A. Tacchella. 2018. Automated Verification of Neural Networks: Advances, Challenges and Perspectives, https://arxiv.org/abs/1805.09938 Marijan, Dusica and Arnaud Gotlieb. Software Testing for Machine Learning. The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20) (2020)

Mirman, M., T. Gehr, and M. Vechev. 2018. Differentiable Abstract Interpretation for Provably Robust Neural Networks. *International Conference on Machine Learning*.

Pullum, Laura L., Brian Taylor, and Marjorie Darrah, *Guidance for the Verification and Validation of Neural Networks*, IEEE Computer Society Press (Wiley), 2007.

Seshia, Sanjit A., Dorsa Sadigh, and S. Shankar Sastry. Towards Verified Artificial Intelligence.

arXiv:1606.08514v4 [cs.AI] 23 Jul 2020.

Taylor, Brian, ed. *Methods and Procedures for the Verification and Validation of Artificial Neural Networks*, Springer-Verlag, 2005.

Xiang, W., P. Musau, A. Wild, D.M. Lopez, N. Hamilton, X. Yang, J. Rosenfeld, and T. Johnson. 2018. Verification for Machine Learning, Autonomy, and Neural Networks Survey. https://arxiv.org/abs/1810.01989

Zhang, J., Li, J. Testing and verification of neural-network-based safety-critical control software: A systematic literature review. *Information and Software Technology* 123, 106296 (2020).

Additional References

Jha, Sumit Kumar, Susmit Jha, Rickard Ewetz, Sunny Raj, Alvaro Velasquez, Laura L. Pullum, and Ananthram Swami. An Extension of Fano's Inequality for Characterizing Model Susceptibility to Membership Inference Attacks. arXiv:2009.08097v1 [cs.LG] 17 Sep 2020.

Sunny Raj, Mesut Ozdag, Steven Fernandes, Sumit Kumar Jha, Laura Pullum, "On the Susceptibility of Deep Neural Networks to Natural Perturbations," *AI Safety 2019* (held in conjunction with IJCAI 2019 - International Joint Conference on Artificial Intelligence), Macao, China, August 2019.

Ak, R., R. Ghosh, G. Shao, H. Reed, Y.-T. Lee, L.L. Pullum. "Verification-Validation and Uncertainty Quantification Methods for Data-Driven Models in Advanced Manufacturing," *ASME Verification and Validation Symposium*, Minneapolis, MN, 2018.

Pullum, L.L., C.A. Steed, S.K. Jha, and A. Ramanathan. "Mathematically Rigorous Verification and Validation of Scientific Machine Learning," *DOE Scientific Machine Learning Workshop*, Bethesda, MD, Jan/Feb 2018.

Ramanathan, A., L.L. Pullum, Zubir Husein, Sunny Raj, Neslisah Totosdagli, Sumanta Pattanaik, and S.K. Jha. 2017. "Adversarial attacks on computer vision algorithms using natural perturbations." In *2017 10th International Conference on Contemporary Computing (IC3)*. Noida, India. August 2017.

Raj, S., L.L. Pullum, A. Ramanathan, and S.K. Jha. 2017. "Work in Progress: Testing Autonomous cyber-physical systems using fuzzing features derived from convolutional neural networks." In *ACM SIGBED International Conference on Embedded Software* (EMSOFT). Seoul, South Korea. October 2017.

Raj, S., L.L. Pullum, A. Ramanathan, and S.K. Jha, "SATYA: Defending against Adversarial Attacks using Statistical Hypothesis Testing," in *10th International Symposium on Foundations and Practice of Security* (FPS 2017), Nancy, France. (Best Paper Award), 2017.

Ramanathan, A., Pullum, L.L., S. Jha, et al. "Integrating Symbolic and Statistical Methods for Testing Intelligent Systems: Applications to Machine Learning and Computer Vision." *IEEE Design, Automation & Test in Europe*(DATE), 2016.

Pullum, L.L., C. Rouff, R. Buskens, X. Cui, E. Vassiv, and M. Hinchey, "Verification of Adaptive Systems," *AIAA Infotech@Aerospace* 2012, April 2012.

Pullum, L.L., and C. Symons, "Failure Analysis of a Complex Learning Framework Incorporating Multi-Modal and Semi-Supervised Learning," In *IEEE Pacific Rim International Symposium on Dependable Computing*(PRDC 2011), 308-313, 2011.

Haglich, P., C. Rouff, and L.L. Pullum, "Detecting Emergent Behaviors with Semi-Boolean Algebra," *Proceedings of AIAA Infotech @ Aerospace*, 2010.

Pullum, L.L., Marjorie A. Darrah, and Brian J. Taylor, "Independent Verification and Validation of Neural Networks – Developing Practitioner Assistance," *Software Tech News*, July 2004.

< Previous Article | Parent Article | Next Article >

SEBoK v. 2.7, released 31 October 2022
