

**2019 ML-MSM Meeting - (October 24-25, 2019)
Poster Abstract Submission Form**

PI(s) Grant:
Institution(s):
Grant Number:
Title of Grant:

Abstract Authors:

Francisco Sahli Costabal^{1,2,3}, Paris Perdikaris⁴, Ellen Kuhl⁵, Daniel E. Hurtado^{1,2,3}

*¹Department of Structural and Geotechnical Engineering, School of Engineering,
Pontificia Universidad Católica de Chile, Santiago, Chile*

*²Institute for Biological and Medical Engineering, Schools of Engineering,
Medicine and Biological Sciences, Pontificia Universidad Católica de Chile,
Santiago, Chile*

³Millennium Nucleus for Cardiovascular Magnetic Resonance

*⁴Department of Mechanical Engineering and Applied Mechanics, University of
Pennsylvania, Philadelphia, Pennsylvania, USA*

*⁵Departments of Mechanical Engineering & Bioengineering, Stanford University,
California, USA*

Abstract Text:

Machine learning techniques typically rely on large datasets to create accurate classifiers. However, there are situations when data is scarce and expensive to acquire. This is the case of studies that rely on state-of-the-art computational models which typically take days to run, thus hindering the potential of machine learning tools. In this work, we present a novel classifier that takes advantage of lower fidelity models and inexpensive approximations to predict the binary output of expensive computer simulations. We postulate an autoregressive model between the different levels of fidelity with Gaussian process priors. We adopt a fully Bayesian treatment for the hyper-parameters and use Markov Chain Monte Carlo samplers. We take advantage of the probabilistic nature of the classifier to implement active learning strategies. We also introduce a sparse approximation to enhance the ability of the multi-fidelity classifier to handle large datasets. We test these multi-fidelity classifiers against their single-fidelity counterpart with synthetic data, showing a median computational cost reduction of 23% for a target accuracy of 90%. In an application to cardiac electrophysiology, the multi-fidelity classifier achieves an F1 score, the harmonic mean of precision and recall, of 99.6% compared to 74.1% of a single-fidelity classifier when both are trained with 50 samples. In general, our results show that the multi-fidelity

classifiers outperform their single-fidelity counterpart in terms of accuracy in all cases. We envision that this new tool will enable researchers to study classification problems that would otherwise be prohibitively expensive. Source code is available at <https://github.com/fsahli/MFclass>