



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
Centro Tecnológico
Colegiado do Curso de Engenharia de Computação

COMUNICADO DE DEFESA DO PROJETO DE GRADUAÇÃO
(Anexo I – Resolução 02/2013-CCEC)

Nome do Estudante: Fernando Guerra Rodrigues Machado

Título: Aprendizado online com seleção de amostras

Data da apresentação: 19/08/2022

Horário: 10h

Local: CT9, Sala 208

Banca Examinadora:

Prof.: Rodolfo da Silva Villaça (Orientador)

Prof. Giovanni Ventorim Comarela (Depto. de Informática)

Prof. Vitor Fontana Zanotelli (PPGI/CT)

Resumo do trabalho: Aprendizagem Online é uma técnica útil para manutenção de modelos de previsões de maneira autônoma. Através dessa técnica conseguimos criar modelos que aprenda continuamente se mantendo atualizados. Porém o volume de amostras geradas pelos sistemas são elevadas e uma grande parcela dessas amostras podem ser já conhecidas pelo modelo atual. Então é desejado criar e manter um dataset reduzido que capture ao máximo o comportamento do sistema, evitando desperdícios de recursos ao armazenar essas amostras e ao atualizar o modelo.

Diante desse cenário, é pertinente o desenvolvimento de técnicas capazes de avaliar se uma amostra é útil para uma futura atualização do modelo, construindo e mantendo um dataset reduzido que capture o comportamento atual do sistema o qual se quer prever. Também é necessário o desenvolvimento de técnicas capazes de identificar quando uma atualização do modelo é necessária, com objetivo de evitar desperdício de poder computacional com atualizações desnecessárias.

Neste trabalho será proposto o *Boxes Sampling*, um algoritmo para selecionar amostras pertinentes e decidir se o modelo precisa ser atualizados. Também será comparado o desempenho do algoritmo proposto com o algoritmo *Reservoir Sampling* que já é conhecido pela comunidade usando o modelo *Random Forest*.