



Imperial College London

Department of Life Sciences

Final Year Project

Supplementary Information

**Feature engineering and selection to improve
autoencoder-based classification of wildtype and
mutant c-Myc molecular dynamics trajectories**

Martina Occhetta

SUPERVISED BY

Dr. Robert O. J. WEINZIERL

June 2022

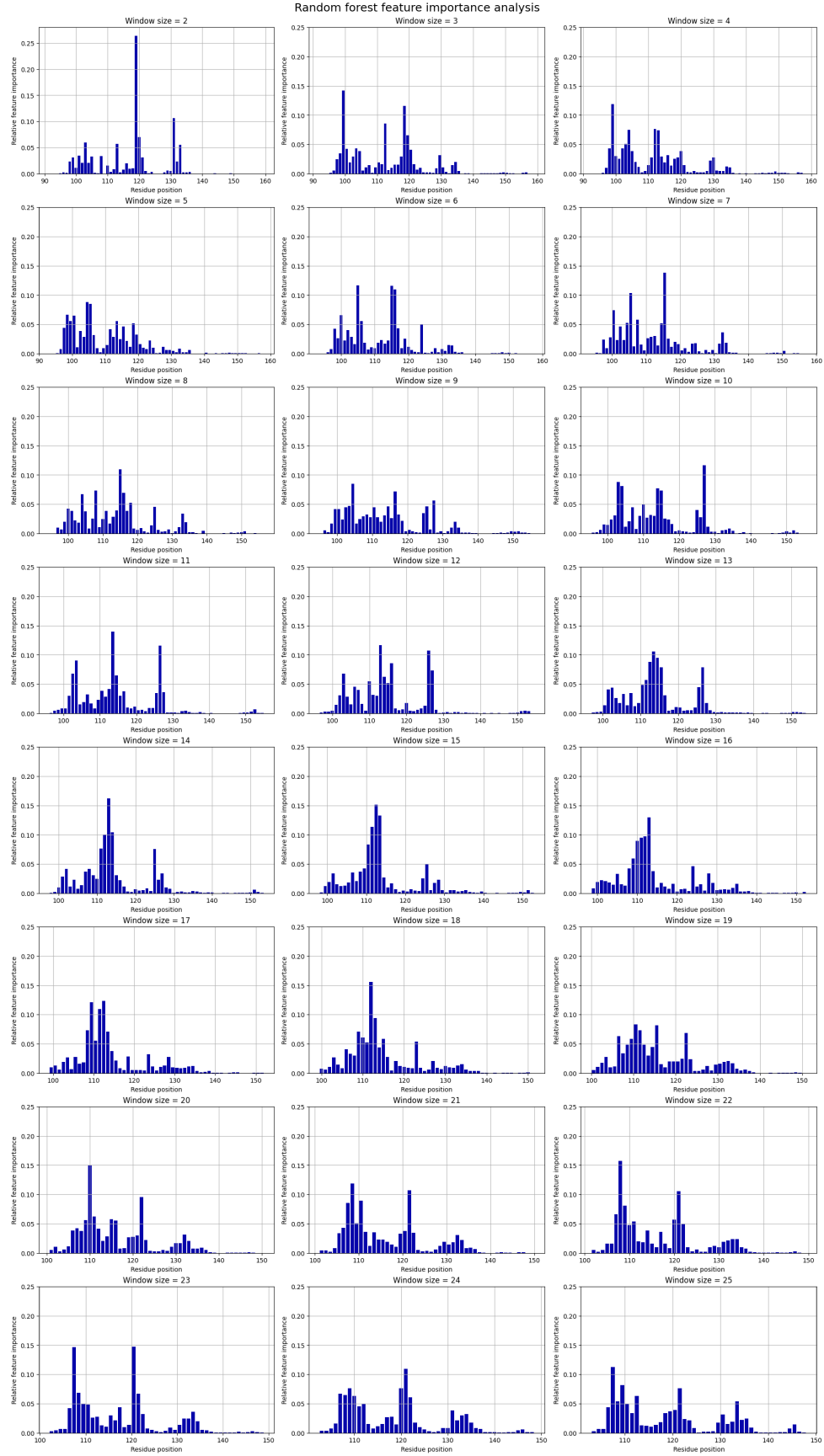


Figure S1: Random forest feature importance analysis for window size 2-25

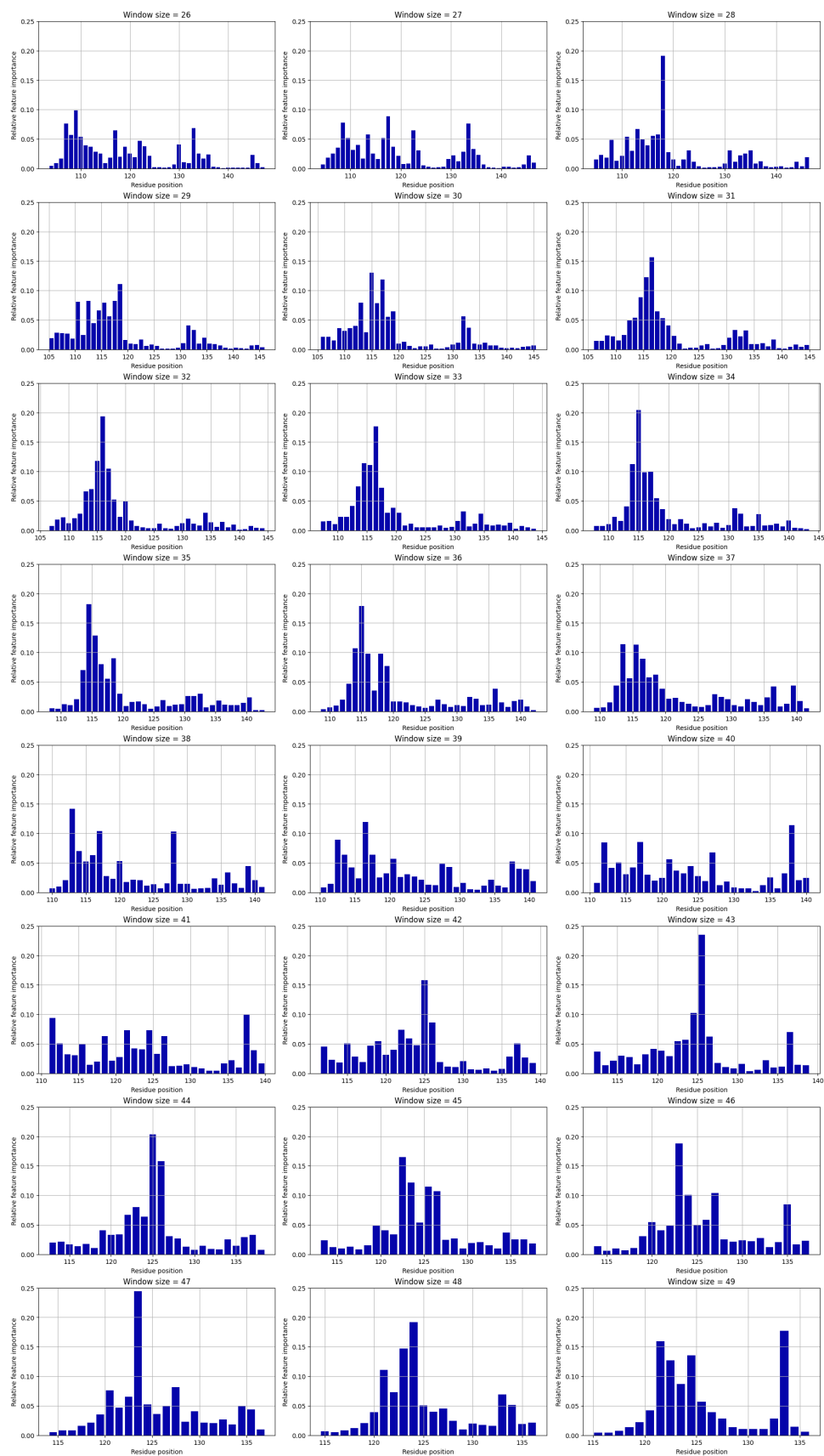


Figure S1: Random forest feature importance analysis for window size 26-50

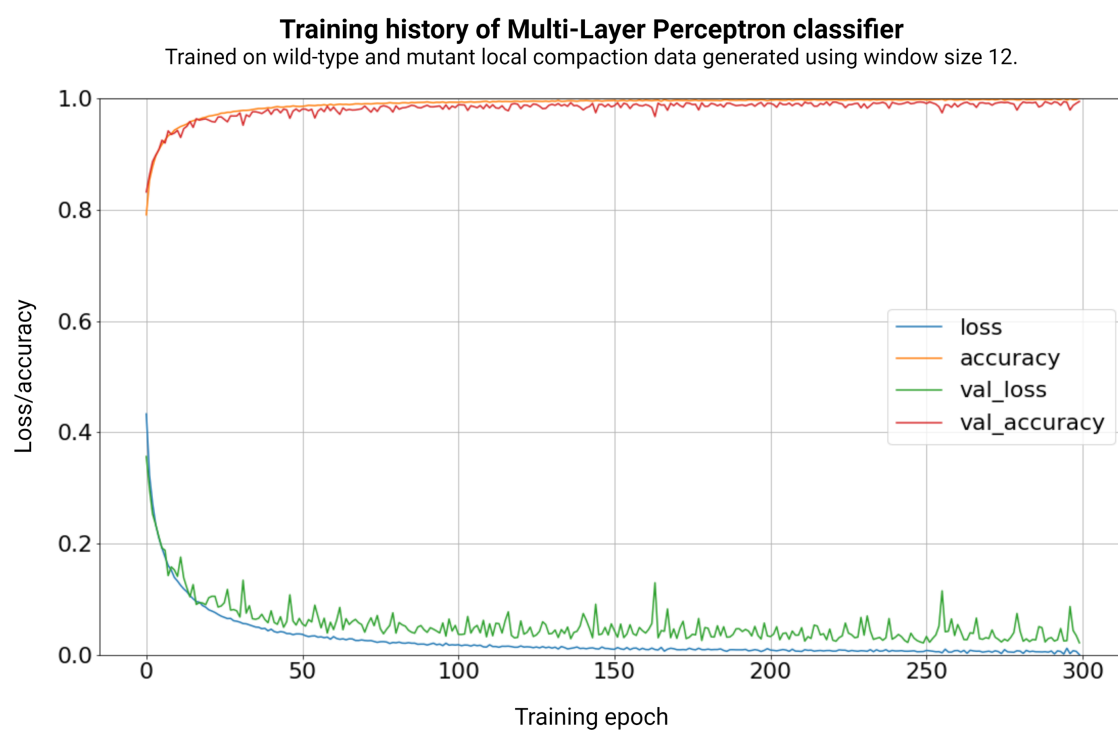


Figure S2: Training history for multi-layer perceptron classifier, trained on wild-type and mutant local compaction data generated using window size = 12.



Figure S3: gzip compression ratios for local compaction data generated using window size 2-25

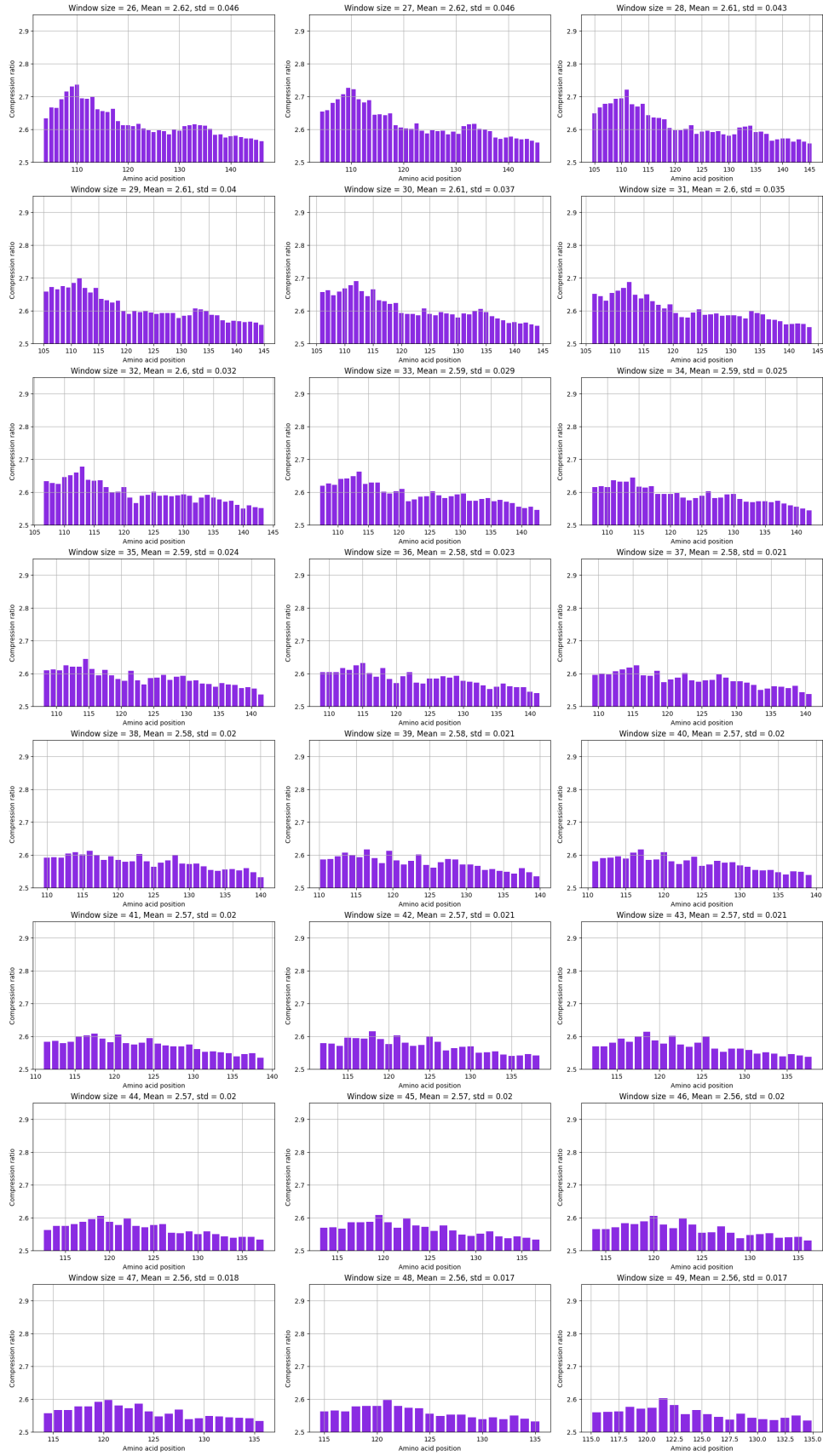


Figure S3: gzip compression ratios for local compaction data generated using window size 26-50