Survey and cross-benchmark comparison of remaining time prediction methods in business process monitoring

THE PAPER EXPLORES PREDICTIVE BUSINESS PROCESS MONITORING, THE MAIN OBJECTIVE IS TO FORECAST TIME-RELATED ASPECTS OF ONGOING PROCESS INSTANCES, LIKE REMAINING TIME AND COMPLETION TIME, USING A SYSTEMATIC REVIEW OF RELEVANT LITERATURE, METHODS CLASSIFICATION, AND REAL-LIFE EVENT LOG EVALUATIONS.
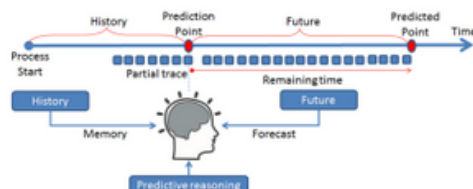

Fig. 2. Overview of predictive process monitoring.

BY FOCUSING ON THE DIVERSE NATURE OF INPUTS AND OUTPUTS, THE ASSESSMENT APPROACH IN THE METHODOLOGICAL FRAMEWORK TACKLES THE FUNDAMENTAL METHODOLOGICAL DISPARITIES. PREFIX BUCKETING IS THE PROCESS OF DIVIDING HISTORICAL LOG PREFIX TRACES INTO BUCKETS FOR SEPARATE PREDICTOR FITTING, AND PREFIX ENCODING IS THE PROCESS OF CONVERTING PREFIXES INTO FIXED-SIZE FEATURE VECTORS FOR MODEL TRAINING. INCLUDING DIVERSE CLASSIFICATION AND REGRESSION METHODS, WITH XGBOOST BEING THE MAIN PREDICTOR, AND EXAMINING A TAXONOMY OF PREDICTIVE MONITORING TECHNIQUES, TAKING INTO ACCOUNT GENERATIVE AND DISCRIMINATIVE APPROACHES.
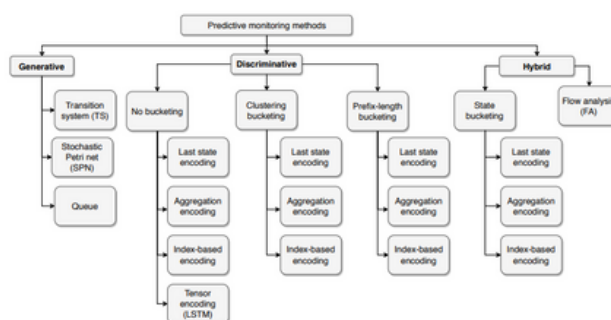

Fig. 5. Taxonomy of methods for predictive monitoring of remaining time.

THE BENCHMARK SECTION EMPLOYS 16 REAL-LIFE EVENT DATASETS, WHICH HAVE BEEN PREPROCESSED TO IMPROVE PREDICTION ACCURACY, AND CONDUCTS EXPERIMENTS BY SPLITTING THE TEMPORAL DATA, MEASURING ACCURACY AND EARLINESS WHILE OPTIMIZING HYPERPARAMETERS. BASED ON THE EVALUATION, IT WAS REVEALED THAT LSTM NETWORKS CONSISTENTLY ACHIEVE SUPERIOR ACCURACY, WITH SIGNIFICANT PERFORMANCE DISPARITIES BETWEEN METHODS. THREATS TO VALIDITY ARE DISCUSSED IN TERMS OF POTENTIAL BIASES IN LITERATURE REVIEW AND EXPERIMENT COMPREHENSIVENESS.

IN CONCLUSION, THE BENCHMARK PROVIDES VALUABLE INSIGHTS INTO PREDICTIVE PROCESS MONITORING APPROACHES THAT PLACE EMPHASIS ON LSTM NETWORKS' EFFECTIVENESS OFFER A STANDARD FRAMEWORK FOR EVALUATION AND COMPARISON, AND ARE SUPPORTED BY PUBLICLY AVAILABLE SOURCE CODE AND DATASETS FOR REPRODUCIBILITY.