Appendix G

Unfolded Results of Perturbation Effectivness Testing on Event Logs

- G.1 BPIC2012
- G.1.1 Decision Tree

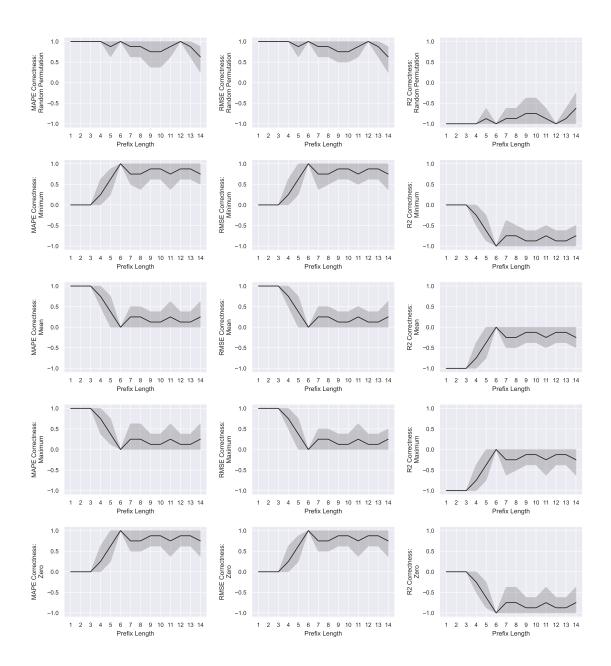


Figure G.1: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with single bucketing and aggregate encoding.

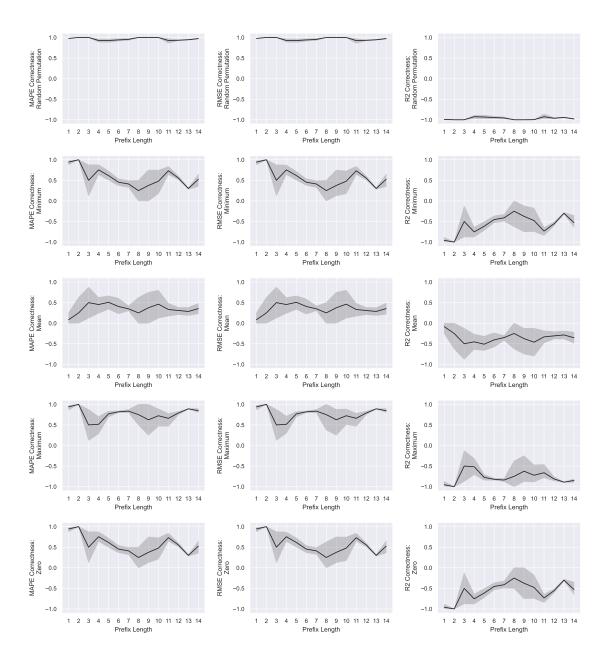


Figure G.2: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with prefix-length bucketing and aggregate encoding.

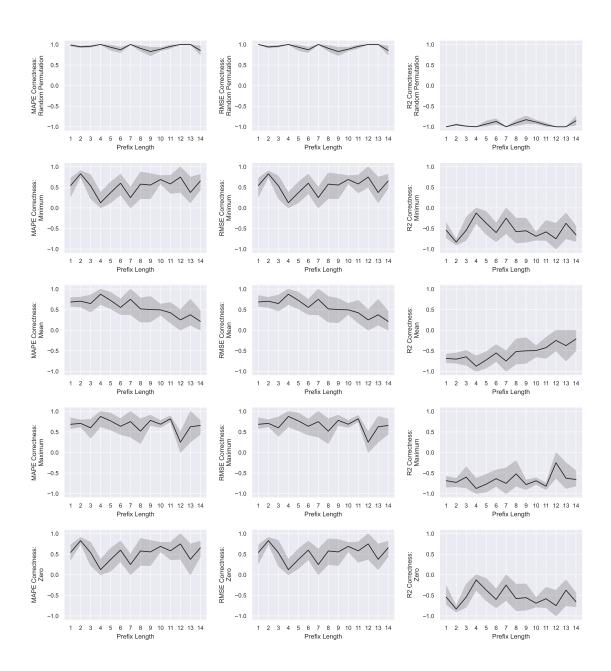


Figure G.3: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with prefix-length bucketing and index-based encoding.

G.1.2 Logistic Regression

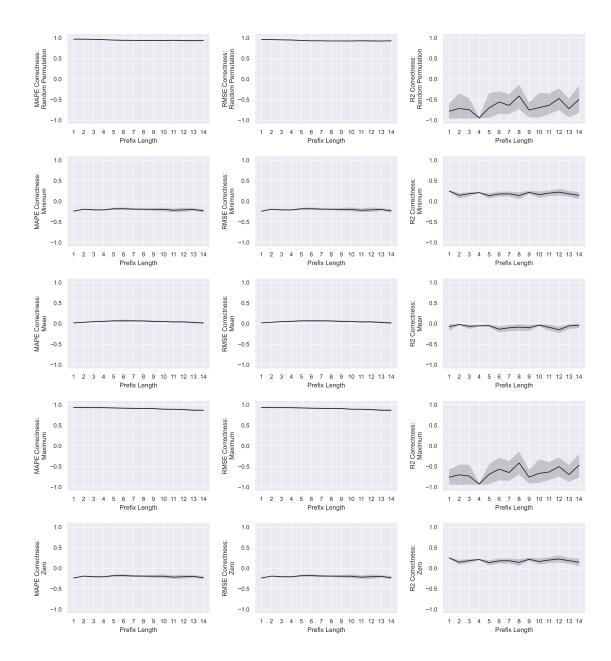


Figure G.4: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with single bucketing and aggregate encoding.

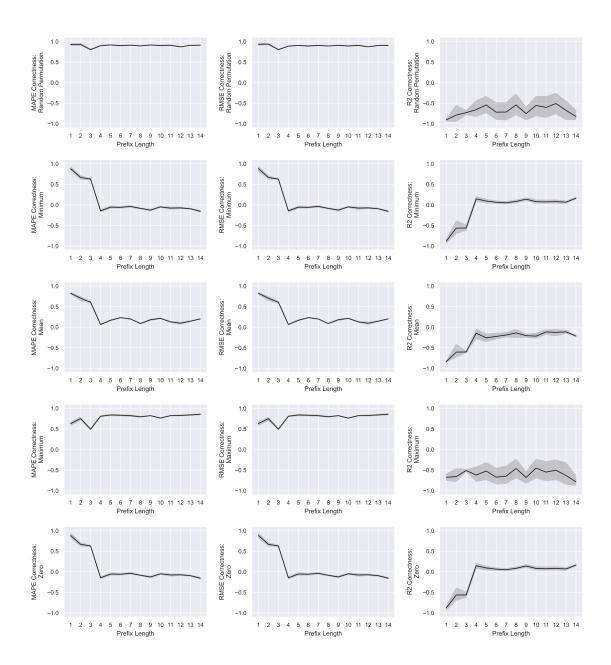


Figure G.5: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with prefix-length bucketing and aggregate encoding.

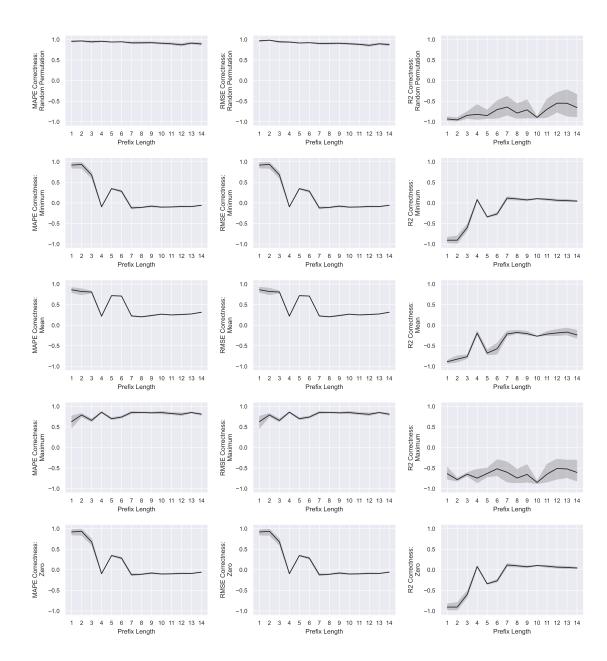


Figure G.6: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with prefix-length bucketing and index-based encoding.

G.1.3 Naïve Bayes

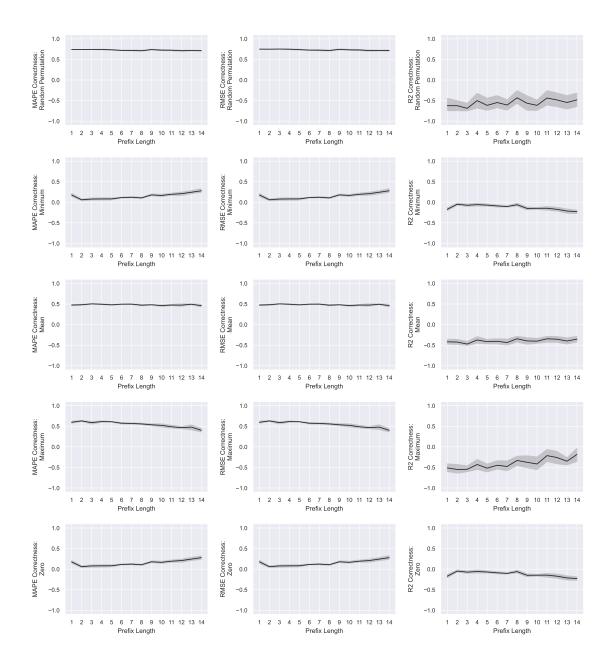


Figure G.7: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with single bucketing and aggregate encoding.

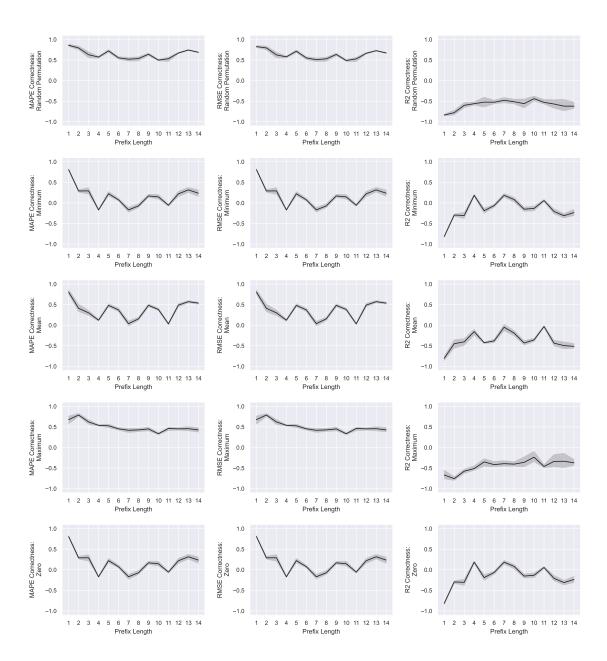


Figure G.8: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with prefix-length bucketing and aggregate encoding.

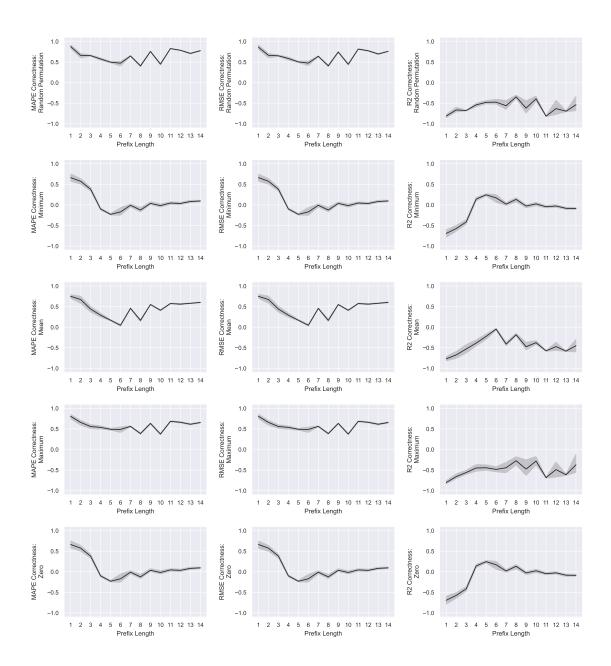


Figure G.9: Effectiveness of perturbation at determining feature importance at each prefix length, using BPIC2012 with prefix-length bucketing and index-based encoding.

G.2 Production

G.2.1 Decision Tree

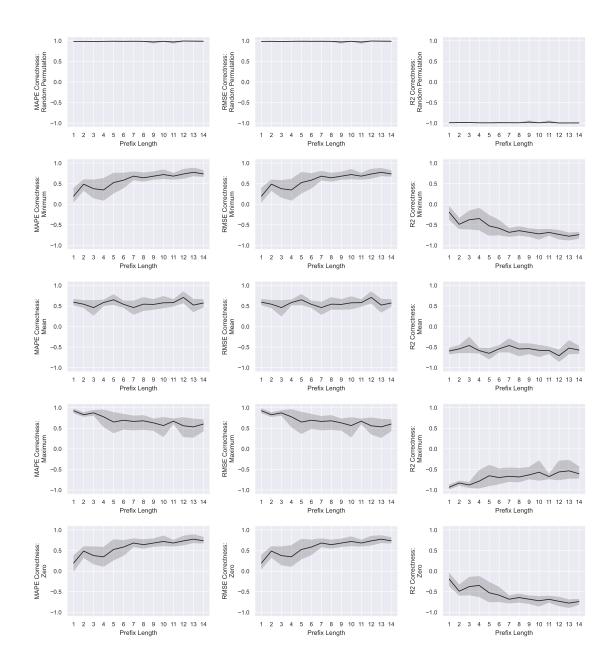


Figure G.10: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with single bucketing and aggregate encoding.

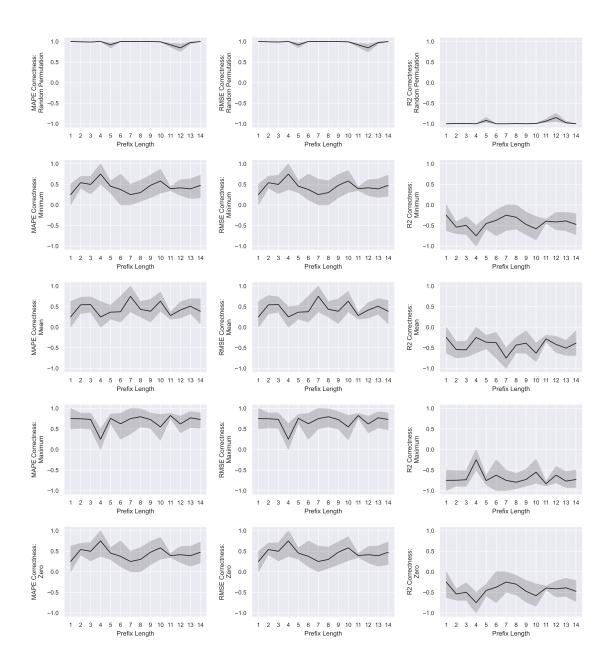


Figure G.11: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with prefix-length bucketing and aggregate encoding.

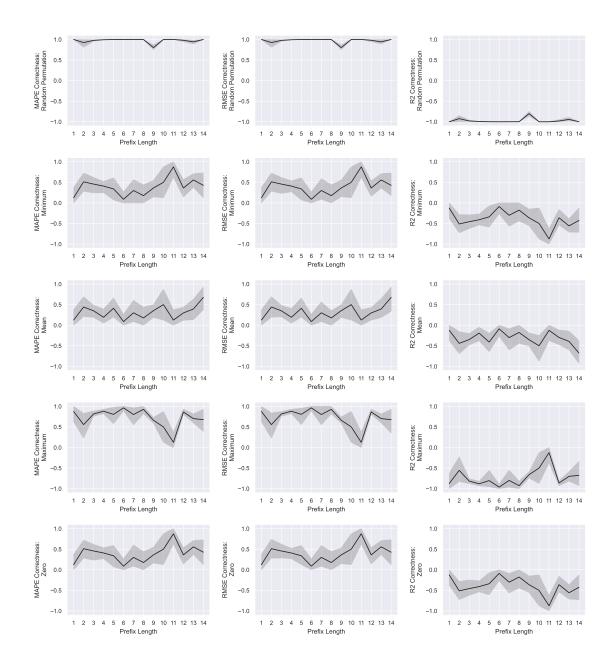


Figure G.12: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with prefix-length bucketing and index-based encoding.

G.2.2 Logistic Regression

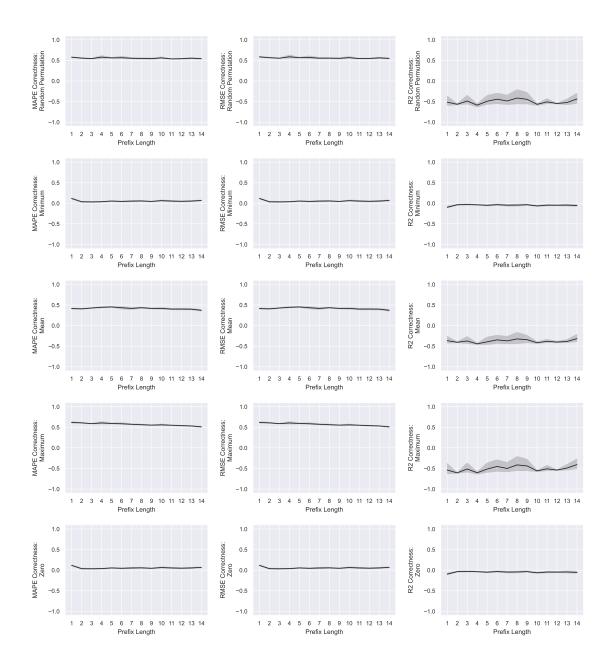


Figure G.13: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with single bucketing and aggregate encoding.

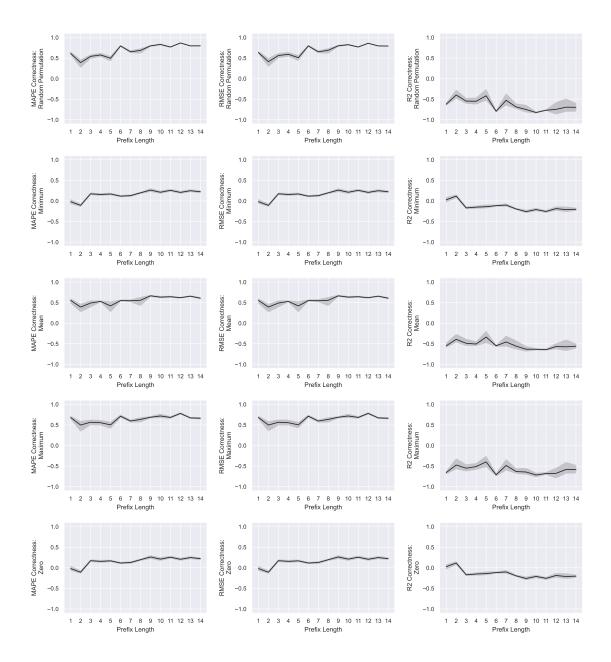


Figure G.14: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with prefix-length bucketing and aggregate encoding.

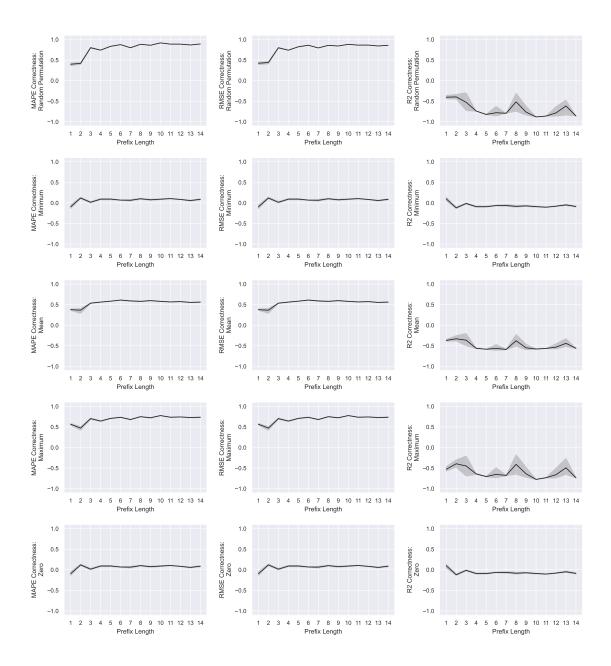


Figure G.15: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with prefix-length bucketing and index-based encoding.

G.2.3 Naïve Bayes

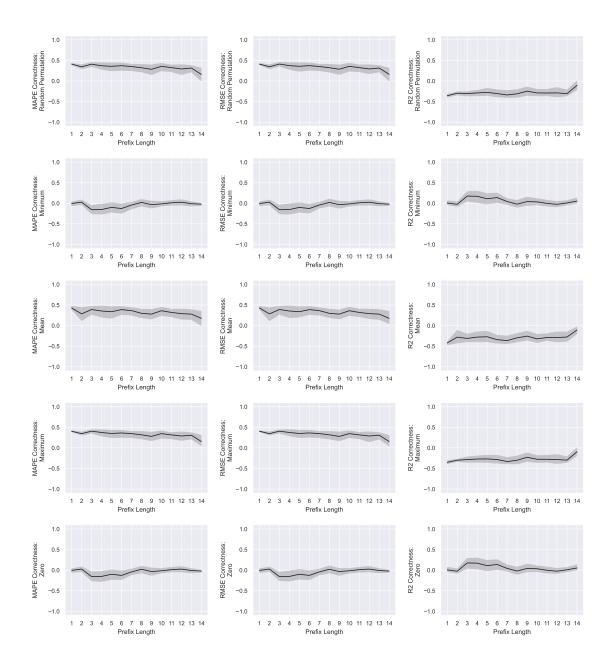


Figure G.16: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with single bucketing and aggregate encoding.

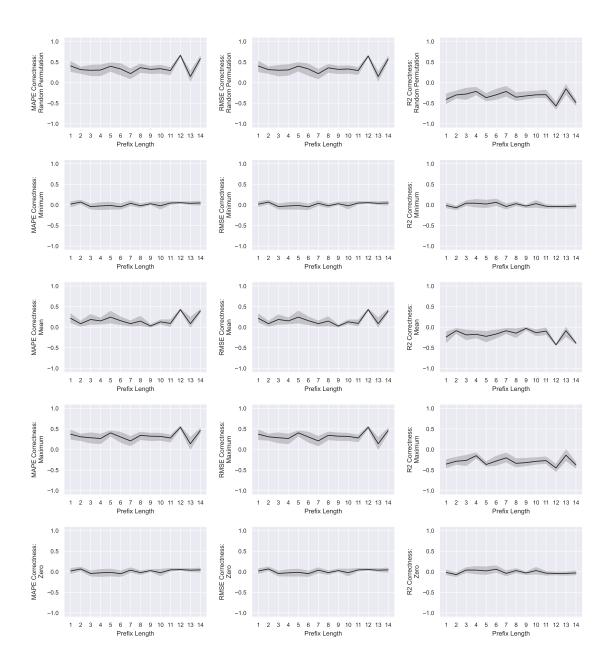


Figure G.17: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with prefix-length bucketing and aggregate encoding.

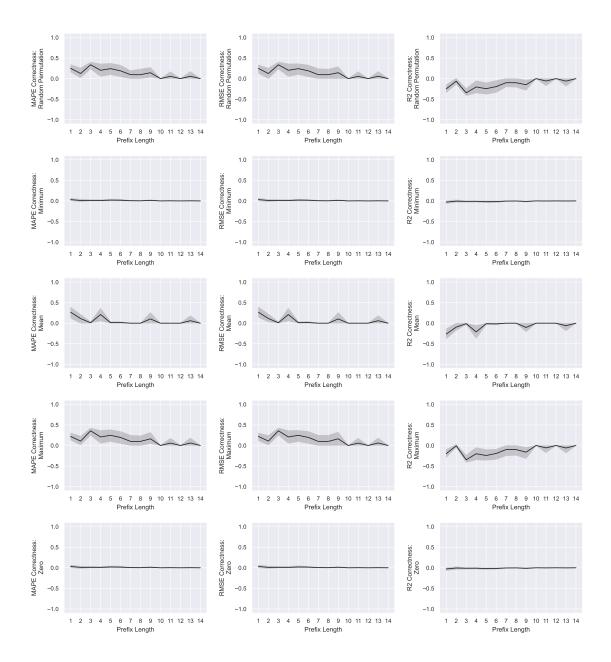


Figure G.18: Effectiveness of perturbation at determining feature importance at each prefix length, using Production with prefix-length bucketing and index-based encoding.

G.3 Sepsis Cases

G.3.1 Decision Tree

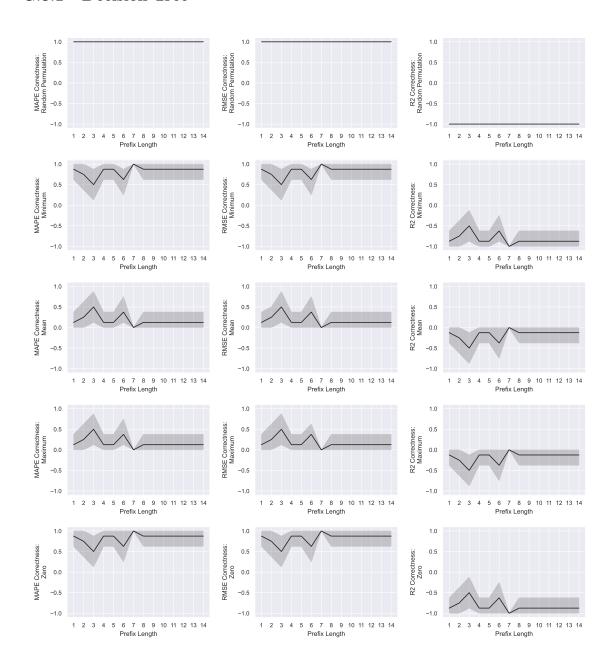


Figure G.19: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with single bucketing and aggregate encoding.

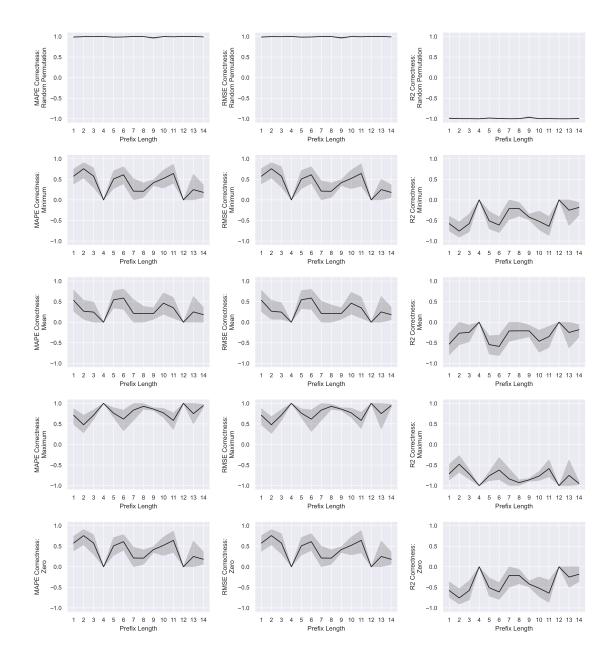


Figure G.20: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with prefix-length bucketing and aggregate encoding.

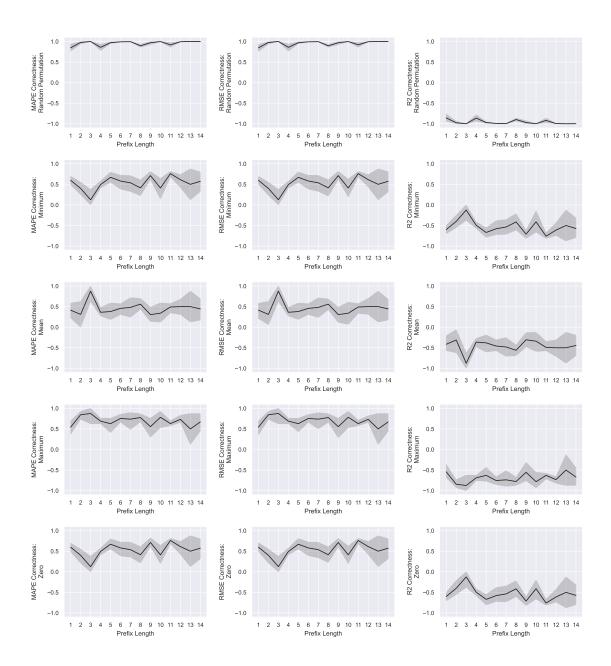


Figure G.21: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with prefix-length bucketing and index-based encoding.

G.3.2 Logistic Regression

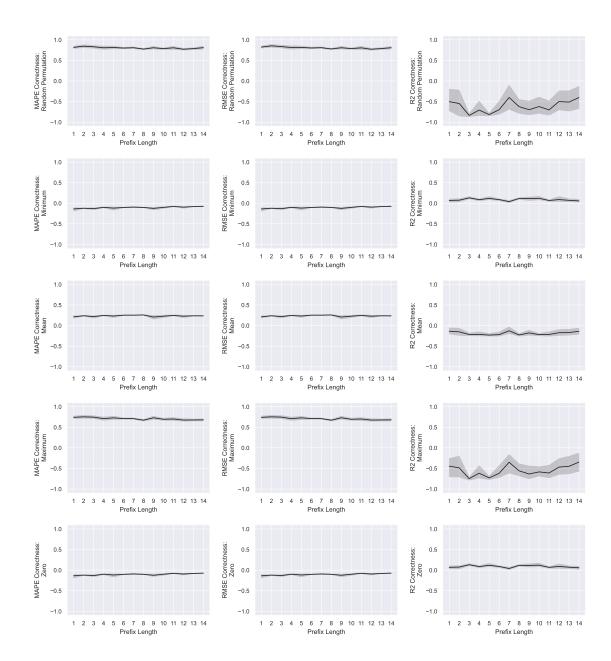


Figure G.22: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with single bucketing and aggregate encoding.

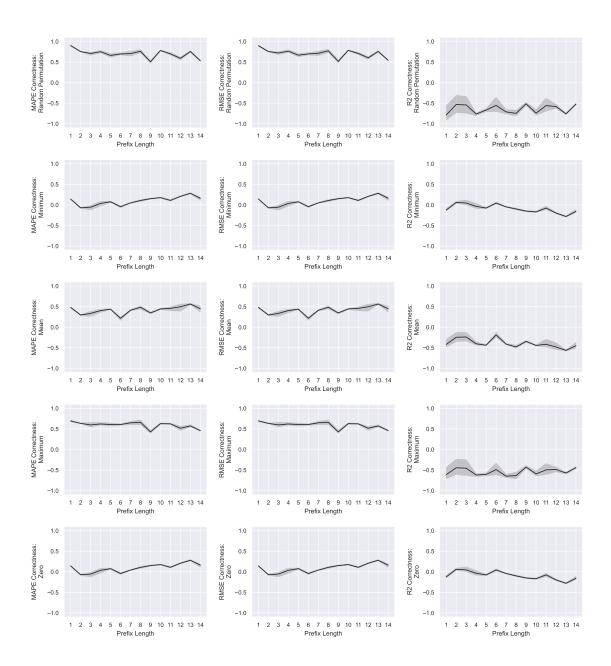


Figure G.23: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with prefix-length bucketing and aggregate encoding.

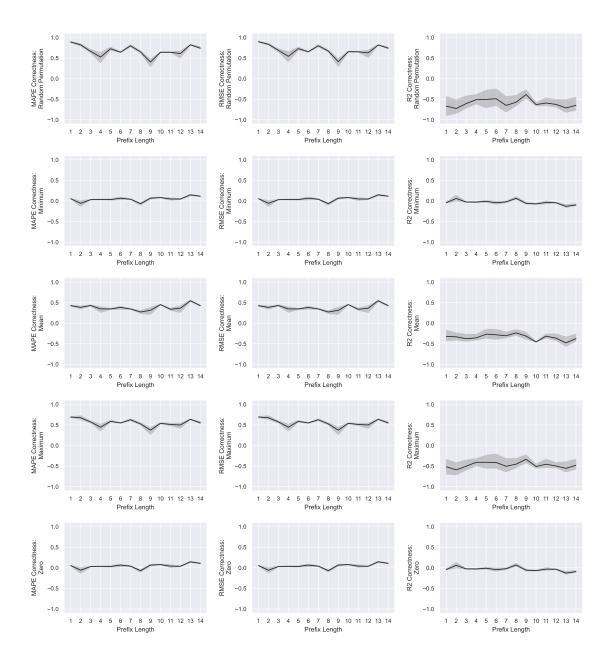


Figure G.24: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with prefix-length bucketing and index-based encoding.

G.3.3 Naïve Bayes

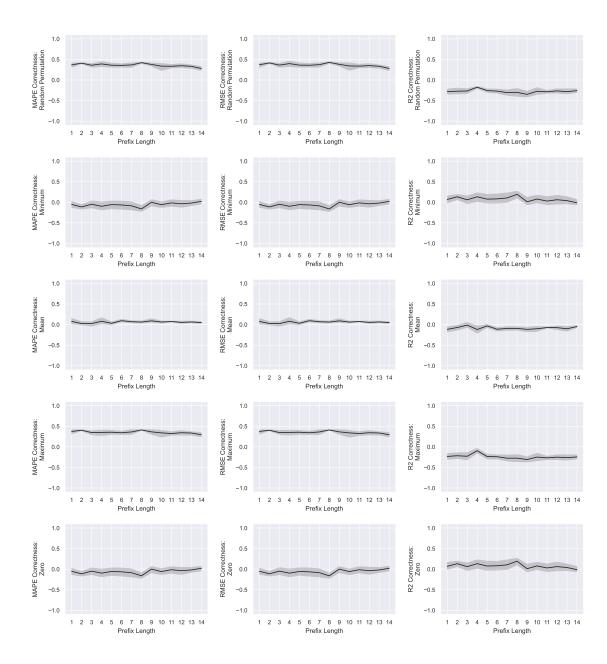


Figure G.25: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with single bucketing and aggregate encoding.

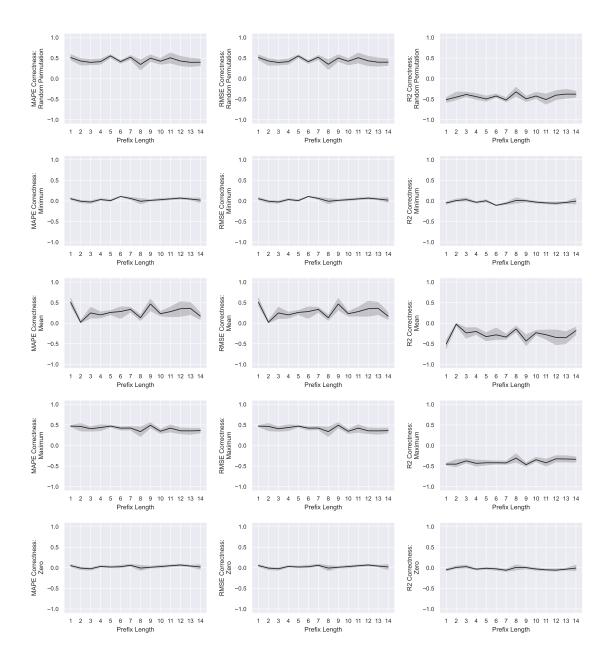


Figure G.26: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with prefix-length bucketing and aggregate encoding.

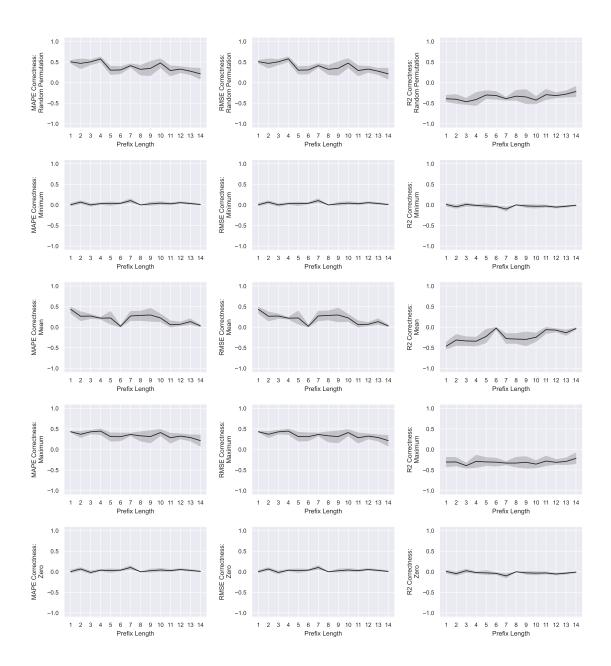


Figure G.27: Effectiveness of perturbation at determining feature importance at each prefix length, using Sepsis Cases with prefix-length bucketing and index-based encoding.