

# Supplementary Material – NCI results

## Paper 146

### 1 Fine-tuning on MIMIC data

Table (1) Accuracy, F1-score and training time on GPU of fine-tuning on MIMIC data with the different optimiser using NN1, and at different learning and drop out rates.

Optimiser	Accuracy (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.902 ± 0.011	0.901 ± 0.012	0.901 ± 0.012	0.904 ± 0.012	0.903 ± 0.011	0.904 ± 0.012
AdamW	0.904 ± 0.013	0.905 ± 0.011	0.904 ± 0.012	0.906 ± 0.012	0.907 ± 0.011	0.908 ± 0.012
SGD	0.806 ± 0.022	0.798 ± 0.027	0.808 ± 0.019	0.871 ± 0.014	0.873 ± 0.014	0.871 ± 0.014

Optimiser	F1-score (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.910 ± 0.010	0.910 ± 0.010	0.910 ± 0.010	0.912 ± 0.010	0.911 ± 0.009	0.912 ± 0.010
AdamW	0.912 ± 0.011	0.913 ± 0.010	0.912 ± 0.011	0.914 ± 0.011	0.915 ± 0.010	0.916 ± 0.010
SGD	0.837 ± 0.018	0.832 ± 0.020	0.839 ± 0.016	0.884 ± 0.012	0.886 ± 0.012	0.884 ± 0.013

Optimiser	Time (Min) (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	63.802 ± 13.346	61.407 ± 11.813	62.738 ± 11.410	28.098 ± 4.932	27.660 ± 5.657	30.205 ± 6.168
AdamW	70.270 ± 19.098	72.570 ± 19.116	70.653 ± 18.160	32.885 ± 8.211	34.152 ± 9.140	31.940 ± 8.840
SGD	97.672 ± 0.940	97.597 ± 0.632	97.725 ± 0.415	97.523 ± 0.539	97.858 ± 0.731	97.688 ± 0.495

Table (2) Accuracy, F1-score and training time on GPU of fine-tuning on MIMIC data with the different optimiser using NN2, and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.947 ± 0.007	0.945 ± 0.013	0.945 ± 0.011	0.942 ± 0.011	0.945 ± 0.012	0.944 ± 0.011
AdamW	0.941 ± 0.009	0.945 ± 0.010	0.944 ± 0.010	0.947 ± 0.010	0.940 ± 0.010	0.942 ± 0.010
SGD	0.833 ± 0.017	0.843 ± 0.020	0.834 ± 0.025	0.922 ± 0.012	0.923 ± 0.014	0.924 ± 0.014

F1-score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.951 ± 0.006	0.949 ± 0.012	0.949 ± 0.010	0.946 ± 0.011	0.949 ± 0.012	0.948 ± 0.010
AdamW	0.945 ± 0.008	0.949 ± 0.009	0.948 ± 0.009	0.951 ± 0.009	0.945 ± 0.009	0.946 ± 0.010
SGD	0.850 ± 0.010	0.859 ± 0.016	0.850 ± 0.020	0.929 ± 0.011	0.929 ± 0.013	0.931 ± 0.013

Time (Min) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	14.695 ± 2.897	14.430 ± 3.161	13.378 ± 2.273	10.672 ± 2.452	10.692 ± 2.094	9.663 ± 2.228
AdamW	8.558 ± 0.426	8.743 ± 0.619	8.850 ± 0.407	7.645 ± 1.080	7.417 ± 0.751	7.463 ± 0.342
SGD	100.977 ± 0.284	101.102 ± 0.633	101.120 ± 1.060	88.945 ± 15.327	89.892 ± 12.977	99.085 ± 4.641

Table (3) Accuracy, F1-score and training time on GPU of fine-tuning on MIMIC data with the different optimiser using NN3, and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.949 ± 0.007	0.948 ± 0.010	0.948 ± 0.011	0.946 ± 0.009	0.946 ± 0.012	0.944 ± 0.010
AdamW	0.946 ± 0.011	0.947 ± 0.008	0.945 ± 0.010	0.946 ± 0.009	0.946 ± 0.011	0.943 ± 0.009
SGD	0.887 ± 0.033	0.891 ± 0.032	0.882 ± 0.036	0.933 ± 0.009	0.935 ± 0.013	0.933 ± 0.010

F1-score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.953 ± 0.007	0.952 ± 0.009	0.952 ± 0.010	0.950 ± 0.009	0.949 ± 0.012	0.948 ± 0.009
AdamW	0.950 ± 0.010	0.952 ± 0.007	0.949 ± 0.009	0.950 ± 0.009	0.950 ± 0.010	0.948 ± 0.008
SGD	0.899 ± 0.027	0.901 ± 0.026	0.895 ± 0.029	0.938 ± 0.008	0.940 ± 0.011	0.938 ± 0.009

Time (Min) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	12.845 ± 3.445	13.223 ± 2.178	12.377 ± 2.736	10.777 ± 1.497	10.902 ± 3.138	11.192 ± 2.510
AdamW	7.690 ± 0.518	7.795 ± 0.617	7.548 ± 0.663	7.007 ± 0.422	7.772 ± 0.810	7.902 ± 1.703
SGD	104.477 ± 0.615	104.742 ± 0.522	104.347 ± 0.521	80.717 ± 18.145	84.183 ± 23.296	83.953 ± 21.729

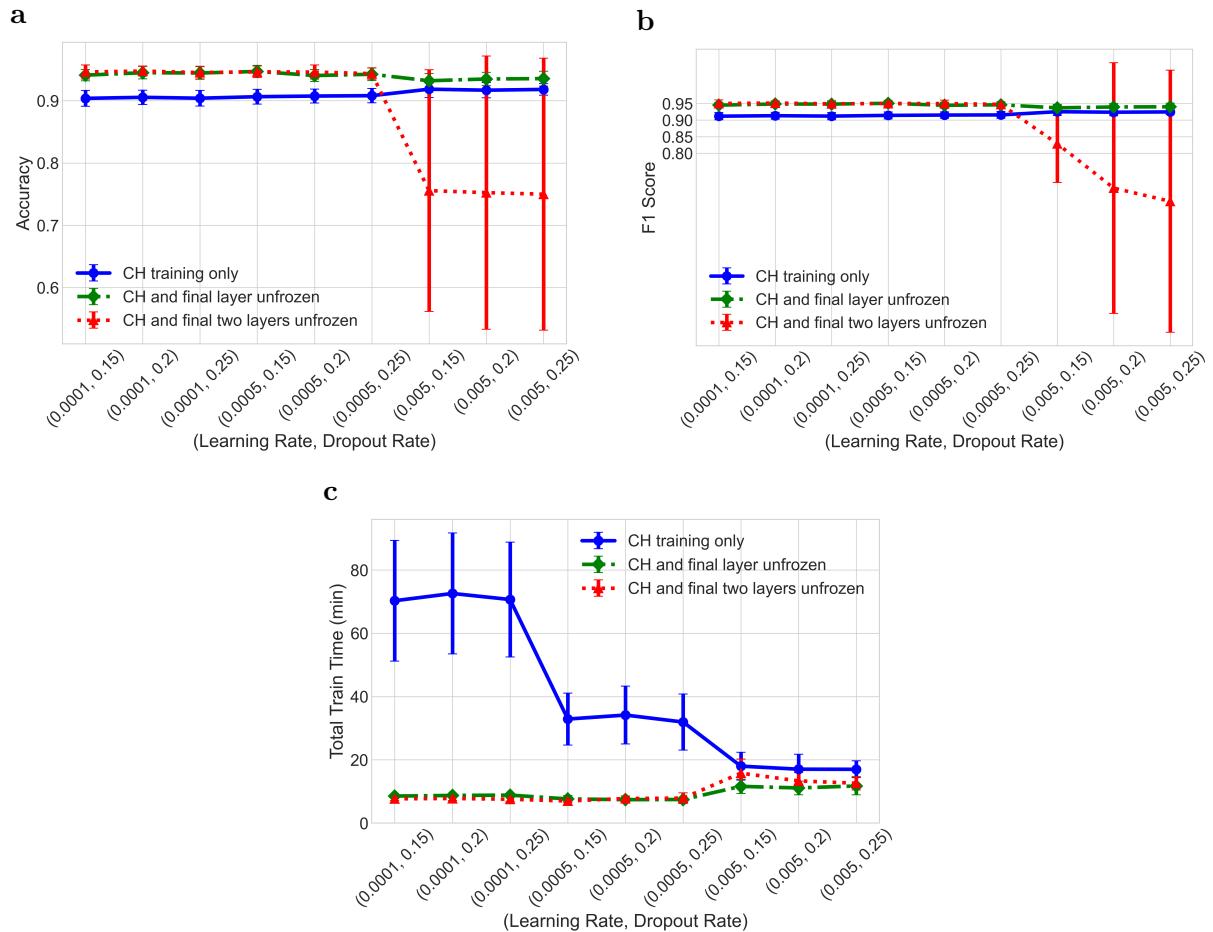


Figure (1) Results of fine-tuning using the AdamW optimiser: (a) Accuracy of the validation data; (b) F-1 score on the validation data; (c) training time.

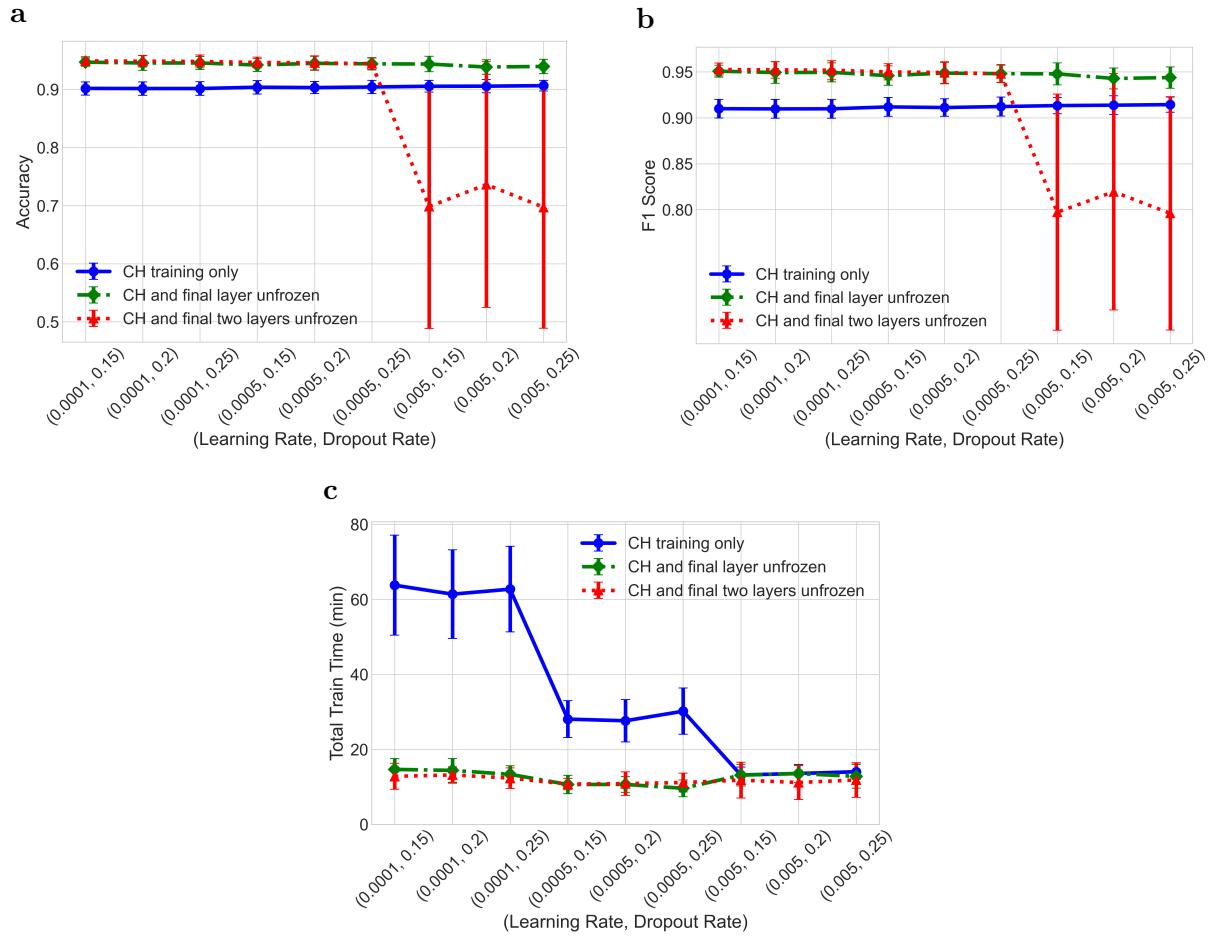


Figure (2) Results of fine-tuning using the Adam optimiser: (a) Accuracy of the validation data; (b) F-1 score on the validation data; (c) training time.

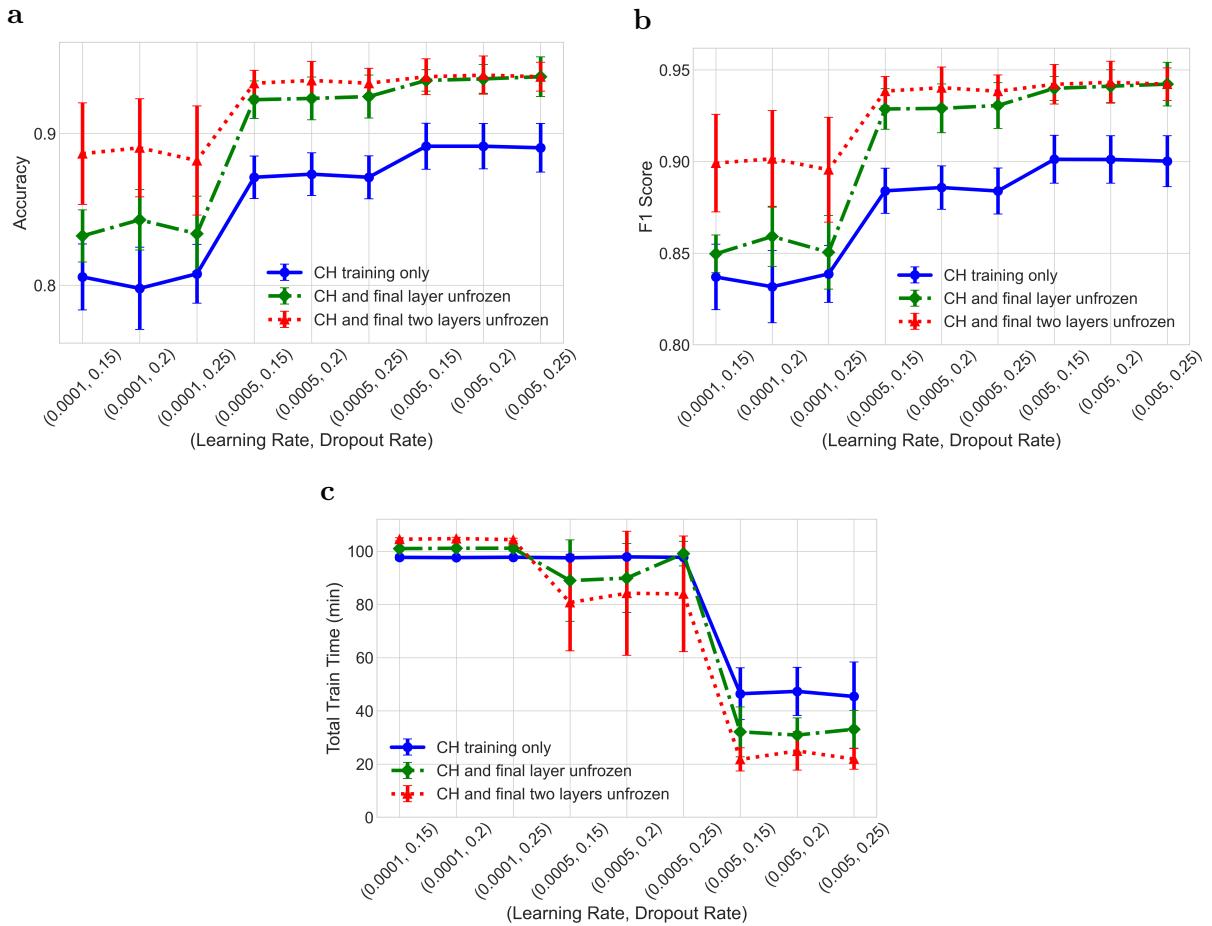


Figure (3) Results of fine-tuning using the SGD optimiser: (a) Accuracy of the validation data; (b) F-1 score on the validation data; (c) training time.

## 2 Prediction with fine-tuned models

Table (4) Prediction results on In-house Two using fine-tuned NN1 with the different optimisers and at different learning and drop out rates.

Optimiser	Accuracy (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.674 ± 0.007	0.670 ± 0.008	0.673 ± 0.009	0.683 ± 0.007	0.683 ± 0.006	0.685 ± 0.007
AdamW	0.681 ± 0.012	0.684 ± 0.014	0.681 ± 0.012	0.694 ± 0.011	0.695 ± 0.010	0.696 ± 0.007
SGD	0.540 ± 0.037	0.514 ± 0.028	0.512 ± 0.025	0.597 ± 0.011	0.598 ± 0.015	0.601 ± 0.010

Optimiser	F1-score (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.675 ± 0.007	0.671 ± 0.008	0.673 ± 0.009	0.684 ± 0.006	0.684 ± 0.006	0.686 ± 0.007
AdamW	0.682 ± 0.012	0.684 ± 0.013	0.682 ± 0.012	0.694 ± 0.011	0.696 ± 0.010	0.696 ± 0.007
SGD	0.527 ± 0.064	0.494 ± 0.039	0.488 ± 0.047	0.600 ± 0.011	0.600 ± 0.015	0.603 ± 0.010

Optimiser	Time (Sec) (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	206.500 ± 27.980	201.900 ± 24.700	201.200 ± 16.250	197.200 ± 18.860	207.600 ± 27.000	201.200 ± 18.740
AdamW	202.400 ± 34.800	199.700 ± 38.490	193.500 ± 17.800	193.300 ± 17.290	192.400 ± 15.730	199.500 ± 27.420
SGD	199.700 ± 17.760	206.500 ± 16.190	236.600 ± 31.800	206.100 ± 20.350	228.400 ± 43.640	207.700 ± 17.880

Table (5) Prediction results on In-house Two using fine-tuned NN2 with the different optimisers and at different learning and drop out rates.

Optimiser	Accuracy (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.819 ± 0.012	0.821 ± 0.016	0.820 ± 0.010	0.827 ± 0.012	0.826 ± 0.019	0.819 ± 0.017
AdamW	0.811 ± 0.015	0.817 ± 0.013	0.817 ± 0.014	0.826 ± 0.016	0.817 ± 0.016	0.819 ± 0.009
SGD	0.563 ± 0.046	0.566 ± 0.045	0.571 ± 0.047	0.728 ± 0.016	0.723 ± 0.016	0.726 ± 0.016

Optimiser	F1-score (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.816 ± 0.012	0.818 ± 0.016	0.817 ± 0.010	0.825 ± 0.013	0.823 ± 0.018	0.816 ± 0.018
AdamW	0.809 ± 0.013	0.816 ± 0.013	0.815 ± 0.012	0.825 ± 0.016	0.816 ± 0.014	0.818 ± 0.008
SGD	0.556 ± 0.053	0.558 ± 0.056	0.562 ± 0.054	0.729 ± 0.015	0.723 ± 0.016	0.726 ± 0.015

Optimiser	Time (Min) (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	184.900 ± 11.230	189.700 ± 13.790	186.500 ± 12.010	197.000 ± 19.060	191.000 ± 17.360	198.300 ± 15.490
AdamW	209.000 ± 22.460	183.400 ± 9.870	199.700 ± 14.010	195.700 ± 14.580	209.200 ± 38.750	198.400 ± 36.530
SGD	218.200 ± 29.640	211.500 ± 22.770	206.500 ± 26.570	189.600 ± 12.640	198.700 ± 30.760	201.500 ± 30.170

Table (6) Prediction results on In-house Two using fine-tuned NN3 using the different optimisers and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.831 ± 0.011	0.833 ± 0.007	0.830 ± 0.012	0.828 ± 0.017	0.831 ± 0.018	0.838 ± 0.012
AdamW	0.817 ± 0.023	0.829 ± 0.012	0.822 ± 0.012	0.829 ± 0.023	0.829 ± 0.008	0.829 ± 0.016
SGD	0.609 ± 0.090	0.629 ± 0.089	0.602 ± 0.083	0.732 ± 0.049	0.741 ± 0.047	0.739 ± 0.047

F1-score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.827 ± 0.010	0.830 ± 0.007	0.826 ± 0.013	0.824 ± 0.018	0.827 ± 0.019	0.834 ± 0.013
AdamW	0.816 ± 0.022	0.828 ± 0.012	0.821 ± 0.012	0.828 ± 0.023	0.828 ± 0.009	0.827 ± 0.017
SGD	0.587 ± 0.112	0.612 ± 0.115	0.581 ± 0.106	0.732 ± 0.049	0.740 ± 0.047	0.739 ± 0.047

Time (Min) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	187.800 ± 13.500	205.500 ± 23.610	185.200 ± 14.120	190.200 ± 12.810	196.500 ± 17.830	196.500 ± 20.210
AdamW	194.500 ± 20.180	192.700 ± 13.750	196.700 ± 10.070	202.900 ± 37.780	196.700 ± 29.340	199.100 ± 30.460
SGD	195.300 ± 13.200	227.700 ± 26.270	218.800 ± 21.350	211.800 ± 23.240	197.900 ± 14.000	195.900 ± 16.970

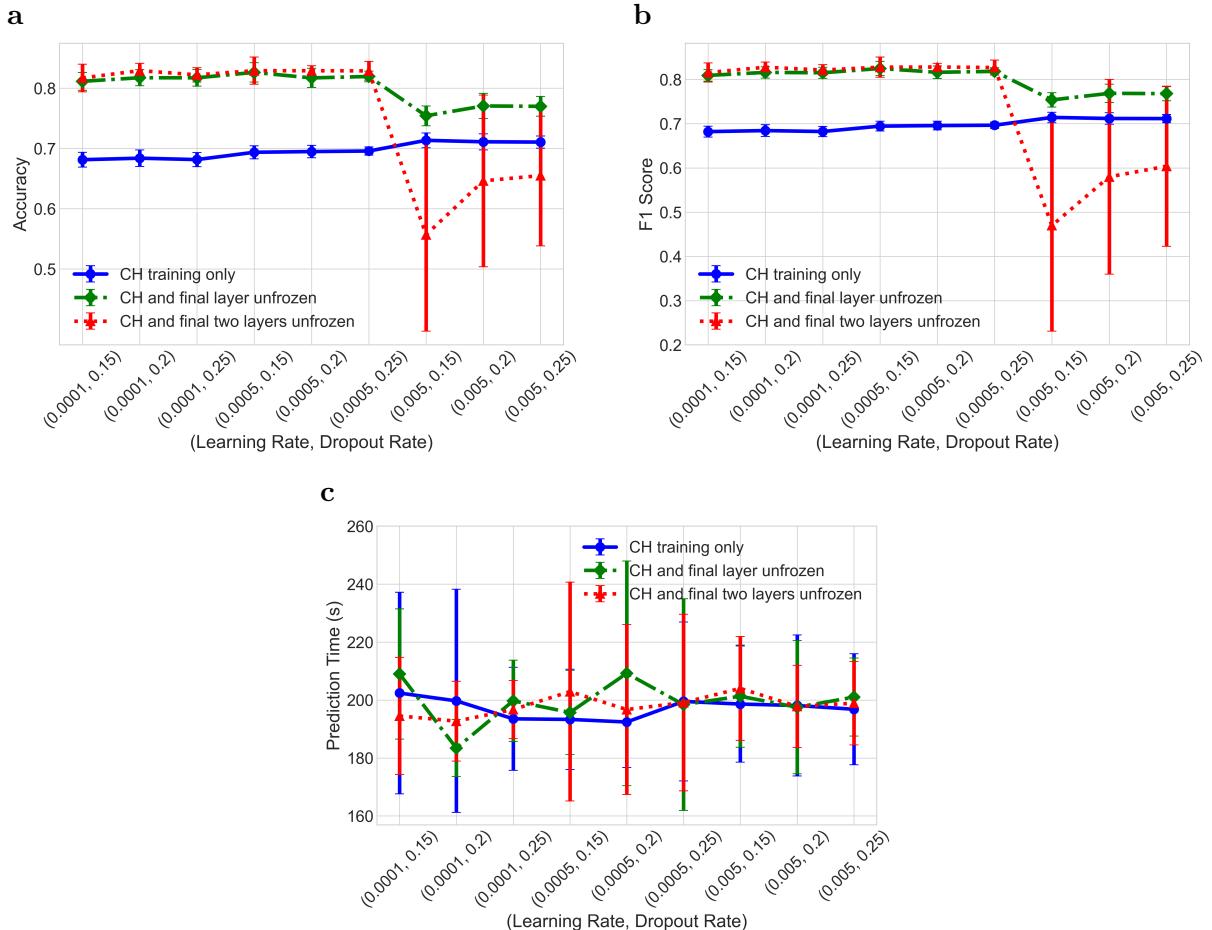


Figure (4) Results of prediction on In-house Two using models fine-tuned with the AdamW optimiser: (a) Prediction accuracy; (b) F-1 score; (c) Prediction on CPU time (s).

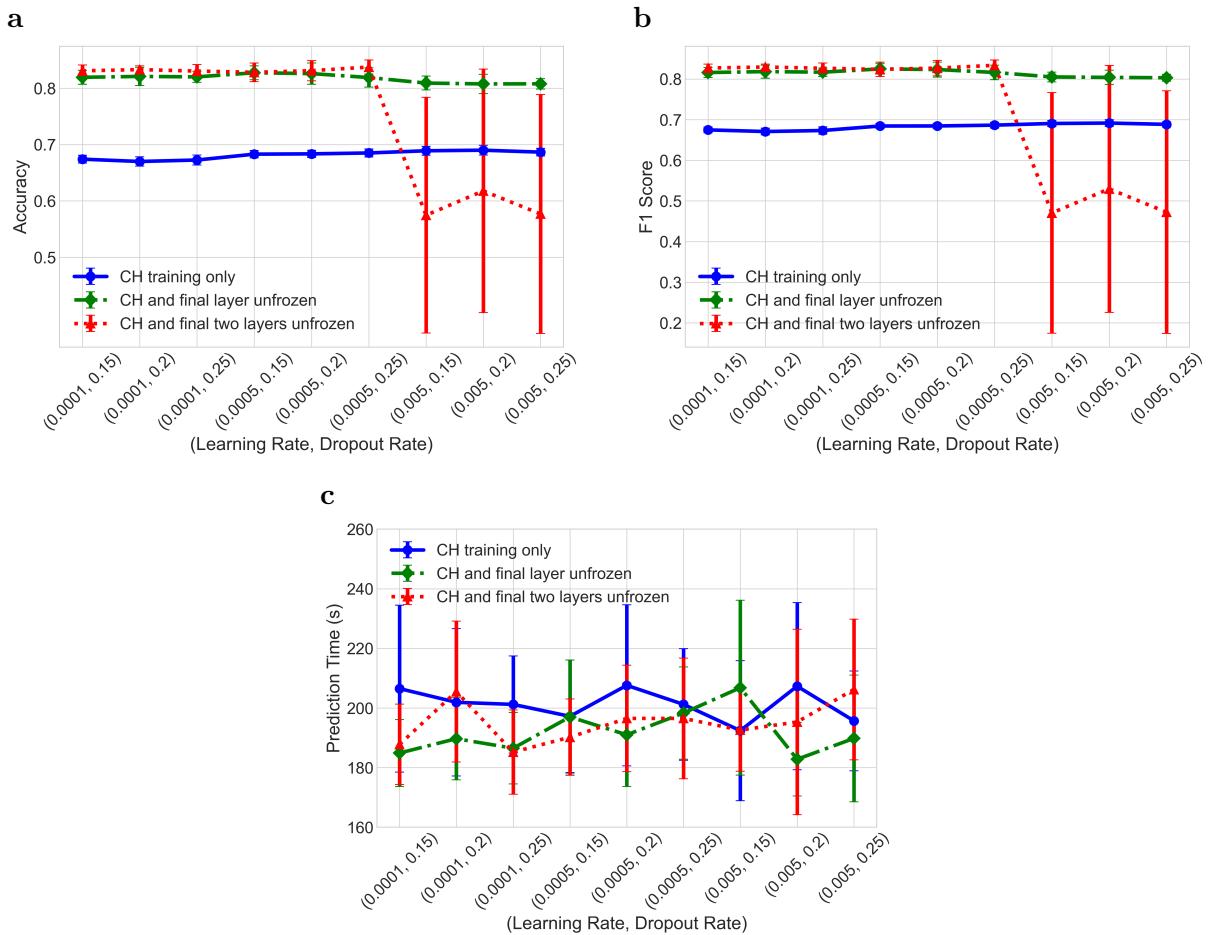


Figure (5) Results of prediction on In-house Two using models fine-tuned with the Adam optimiser: (a) Prediction accuracy; (b) F-1 score; (c) Prediction on CPU time (s).

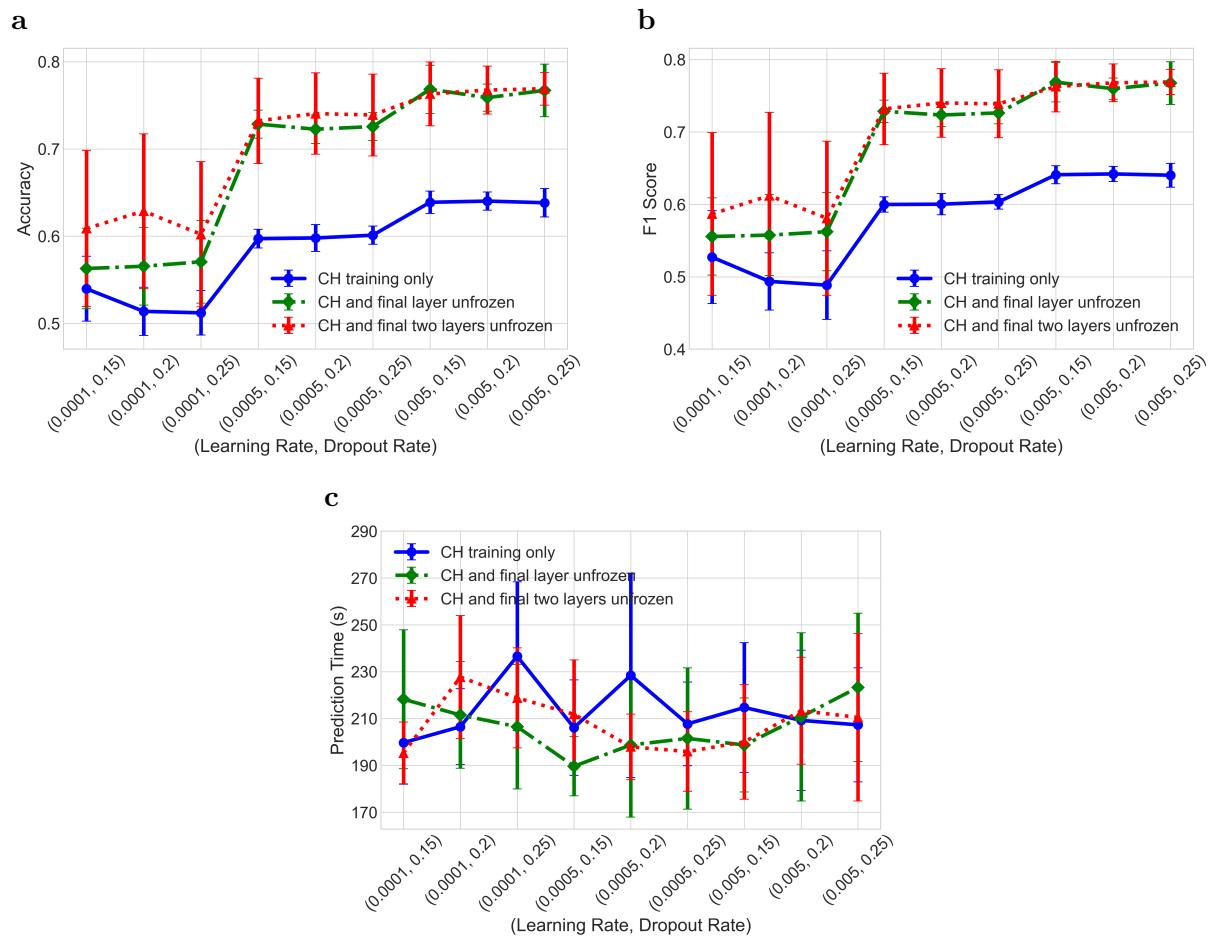


Figure (6) Results of prediction on In-house Two using models fine-tuned with the SGD optimiser: (a) Prediction accuracy; (b) F-1 score; (c) Prediction on CPU time (s).

### 3 Further fine-tuning on In-House One Data

Table (7) Accuracy, F1-score and training time on CPU of further fine-tuning (domain adaptation) on In-house One data with the different optimiser using NN1, and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.683 ± 0.046	0.690 ± 0.040	0.682 ± 0.044	0.728 ± 0.030	0.732 ± 0.034	0.721 ± 0.031
AdamW	0.682 ± 0.035	0.684 ± 0.043	0.686 ± 0.036	0.741 ± 0.031	0.737 ± 0.030	0.740 ± 0.037
F1 Score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.610 ± 0.053	0.617 ± 0.046	0.609 ± 0.050	0.656 ± 0.044	0.662 ± 0.048	0.652 ± 0.040
AdamW	0.617 ± 0.034	0.621 ± 0.043	0.622 ± 0.040	0.685 ± 0.041	0.680 ± 0.040	0.686 ± 0.042
Time (min) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	318.072 ± 309.069	413.718 ± 399.177	405.935 ± 432.426	275.267 ± 95.007	276.672 ± 92.814	254.755 ± 98.248
AdamW	434.120 ± 456.440	343.088 ± 350.208	359.795 ± 314.374	410.525 ± 306.164	295.127 ± 82.916	301.505 ± 120.091

Table (8) Accuracy, F1-score and training time on CPU of further fine-tuning (domain adaptation) on In-house One data with the different optimiser using NN2, and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.936 ± 0.009	0.931 ± 0.006	0.935 ± 0.012	0.941 ± 0.008	0.940 ± 0.013	0.941 ± 0.010
AdamW	0.923 ± 0.012	0.923 ± 0.013	0.921 ± 0.016	0.935 ± 0.013	0.934 ± 0.010	0.933 ± 0.013
F1 Score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.930 ± 0.010	0.925 ± 0.007	0.929 ± 0.012	0.935 ± 0.009	0.934 ± 0.014	0.935 ± 0.010
AdamW	0.917 ± 0.012	0.917 ± 0.014	0.914 ± 0.017	0.929 ± 0.014	0.929 ± 0.010	0.927 ± 0.014
Time (min) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	135.895 ± 31.279	126.288 ± 18.203	141.438 ± 30.566	113.547 ± 29.327	98.768 ± 22.812	97.407 ± 15.982
AdamW	99.932 ± 20.437	99.948 ± 14.998	125.565 ± 14.869	98.928 ± 34.014	91.208 ± 26.534	104.963 ± 24.340

Table (9) Accuracy, F1-score and training time on CPU of further fine-tuning (domain adaptation) on In-house One data with the different optimiser using NN3, and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.944 $\pm$ 0.015	0.942 $\pm$ 0.014	0.942 $\pm$ 0.015	0.941 $\pm$ 0.010	0.938 $\pm$ 0.010	0.939 $\pm$ 0.010
AdamW	0.937 $\pm$ 0.009	0.940 $\pm$ 0.014	0.938 $\pm$ 0.013	0.937 $\pm$ 0.007	0.940 $\pm$ 0.013	0.940 $\pm$ 0.015

F1 Score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.939 $\pm$ 0.016	0.937 $\pm$ 0.014	0.936 $\pm$ 0.016	0.934 $\pm$ 0.011	0.931 $\pm$ 0.011	0.932 $\pm$ 0.012
AdamW	0.931 $\pm$ 0.009	0.934 $\pm$ 0.014	0.932 $\pm$ 0.014	0.931 $\pm$ 0.008	0.934 $\pm$ 0.014	0.933 $\pm$ 0.015

Time (min) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	128.622 $\pm$ 39.316	109.002 $\pm$ 27.680	125.325 $\pm$ 31.736	105.012 $\pm$ 23.066	98.093 $\pm$ 16.029	96.058 $\pm$ 17.986
AdamW	87.885 $\pm$ 14.763	100.350 $\pm$ 25.593	105.780 $\pm$ 17.340	97.383 $\pm$ 11.519	102.645 $\pm$ 31.401	108.853 $\pm$ 21.445

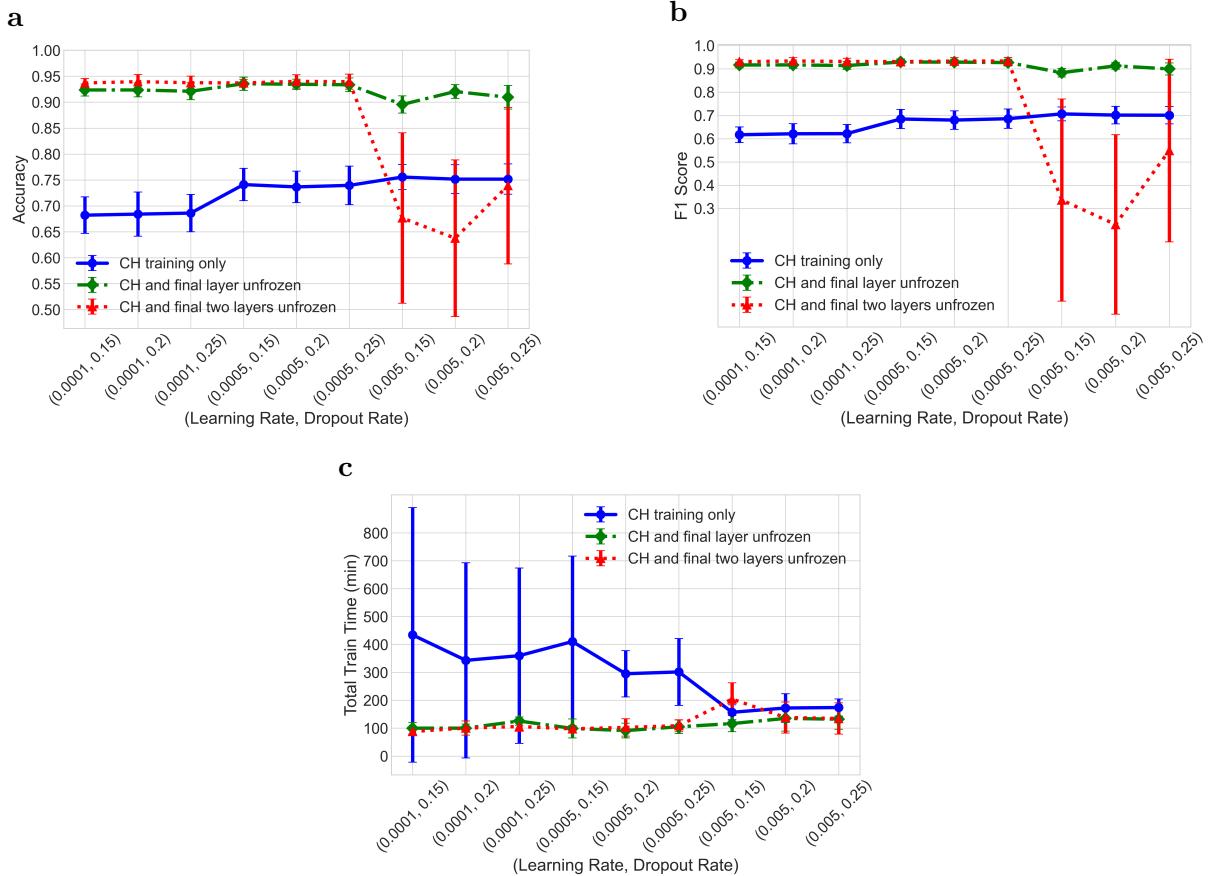


Figure (7) Results of further fine-tuning on In-house One data with the AdamW optimiser: (a) Prediction accuracy; (b) F-1 score; (c) Training time on CPU (min).

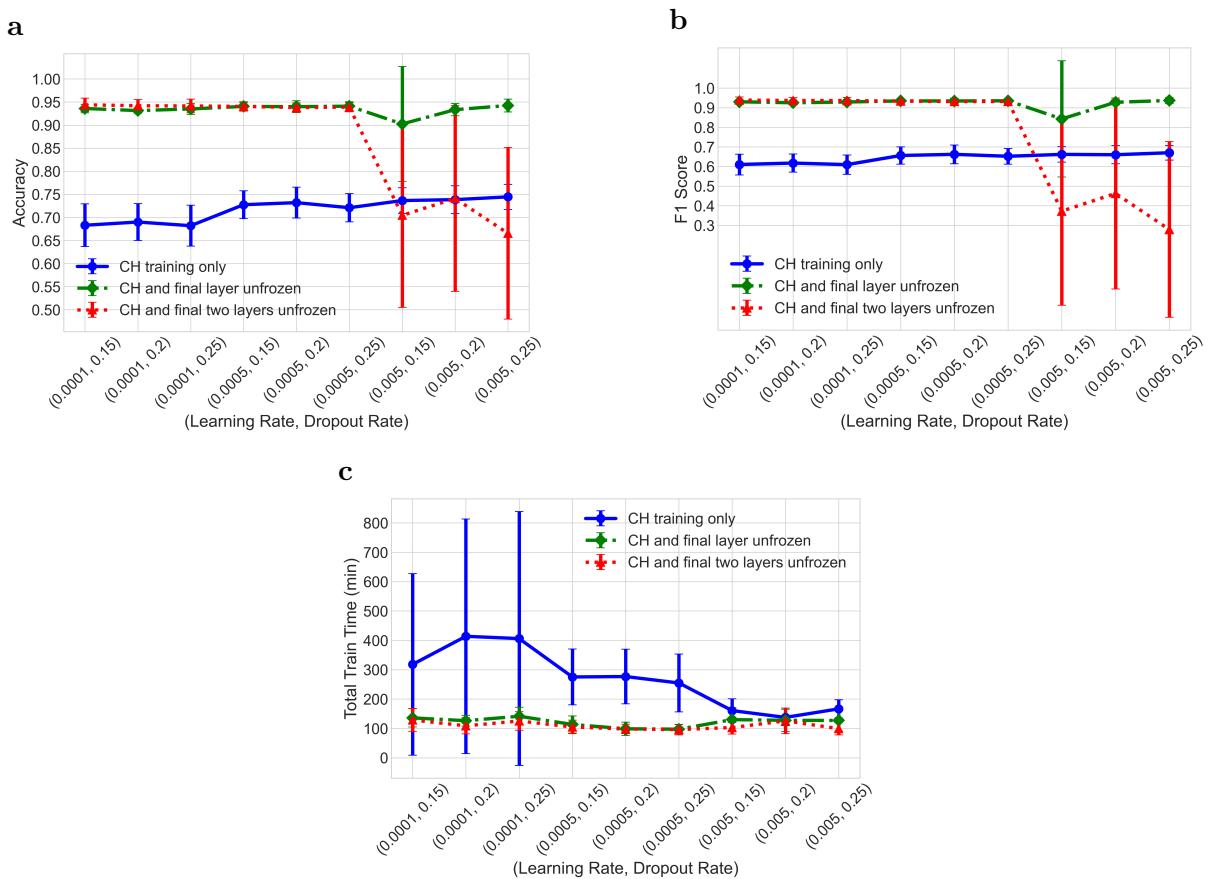


Figure (8) Results of further fine-tuning on In-house One data with the Adam optimiser:  
(a) Prediction accuracy; (b) F-1 score; (c) Training time on CPU (min).

## 4 Prediction with domain adapted models

Table (10) Accuracy, F1-score and training time on CPU of further fine-tuning (domain adaptation) on In-house One data with the different optimiser using NN1, and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.931 ± 0.004	0.931 ± 0.005	0.931 ± 0.005	0.936 ± 0.005	0.933 ± 0.004	0.933 ± 0.004
AdamW	0.922 ± 0.006	0.923 ± 0.006	0.924 ± 0.008	0.929 ± 0.006	0.926 ± 0.006	0.930 ± 0.005
F1 Score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.920 ± 0.004	0.919 ± 0.005	0.919 ± 0.006	0.925 ± 0.006	0.921 ± 0.004	0.921 ± 0.004
AdamW	0.909 ± 0.006	0.910 ± 0.006	0.912 ± 0.008	0.917 ± 0.006	0.915 ± 0.007	0.919 ± 0.005
Time (Sec) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	202.300 ± 33.010	212.600 ± 26.750	195.900 ± 22.750	217.100 ± 29.440	201.700 ± 13.710	201.900 ± 21.240
AdamW	226.800 ± 39.670	195.900 ± 19.920	200.000 ± 16.750	211.300 ± 25.640	217.300 ± 42.830	218.600 ± 24.650

Table (11) Accuracy, F1-score and training time on CPU of further fine-tuning (domain adaptation) on In-house One data with the different optimiser using NN2, and at different learning and drop out rates.

Accuracy (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.931 ± 0.004	0.931 ± 0.005	0.931 ± 0.005	0.936 ± 0.005	0.933 ± 0.004	0.933 ± 0.004
AdamW	0.922 ± 0.006	0.923 ± 0.006	0.924 ± 0.008	0.929 ± 0.006	0.926 ± 0.006	0.930 ± 0.005
F1 Score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.920 ± 0.004	0.919 ± 0.005	0.919 ± 0.006	0.925 ± 0.006	0.921 ± 0.004	0.921 ± 0.004
AdamW	0.909 ± 0.006	0.910 ± 0.006	0.912 ± 0.008	0.917 ± 0.006	0.915 ± 0.007	0.919 ± 0.005
Time (Sec) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	202.300 ± 33.010	212.600 ± 26.750	195.900 ± 22.750	217.100 ± 29.440	201.700 ± 13.710	201.900 ± 21.240
AdamW	226.800 ± 39.670	195.900 ± 19.920	200.000 ± 16.750	211.300 ± 25.640	217.300 ± 42.830	218.600 ± 24.650

Table (12) Accuracy, F1-score and training time on CPU of further fine-tuning (domain adaptation) on In-house One data with the different optimiser using NN3, and at different learning and drop out rates.

Optimiser	Accuracy (Learning rate, Drop out rate)					
	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.931 ± 0.004	0.931 ± 0.005	0.931 ± 0.005	0.936 ± 0.005	0.933 ± 0.004	0.933 ± 0.004
AdamW	0.922 ± 0.006	0.923 ± 0.006	0.924 ± 0.008	0.929 ± 0.006	0.926 ± 0.006	0.930 ± 0.005
F1 Score (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	0.920 ± 0.004	0.919 ± 0.005	0.919 ± 0.006	0.925 ± 0.006	0.921 ± 0.004	0.921 ± 0.004
AdamW	0.909 ± 0.006	0.910 ± 0.006	0.912 ± 0.008	0.917 ± 0.006	0.915 ± 0.007	0.919 ± 0.005
Time (Sec) (Learning rate, Drop out rate)						
Optimiser	(0.0001, 0.15)	(0.0001, 0.2)	(0.0001, 0.25)	(0.0005, 0.15)	(0.0005, 0.2)	(0.0005, 0.25)
Adam	202.300 ± 33.010	212.600 ± 26.750	195.900 ± 22.750	217.100 ± 29.440	201.700 ± 13.710	201.900 ± 21.240
AdamW	226.800 ± 39.670	195.900 ± 19.920	200.000 ± 16.750	211.300 ± 25.640	217.300 ± 42.830	218.600 ± 24.650

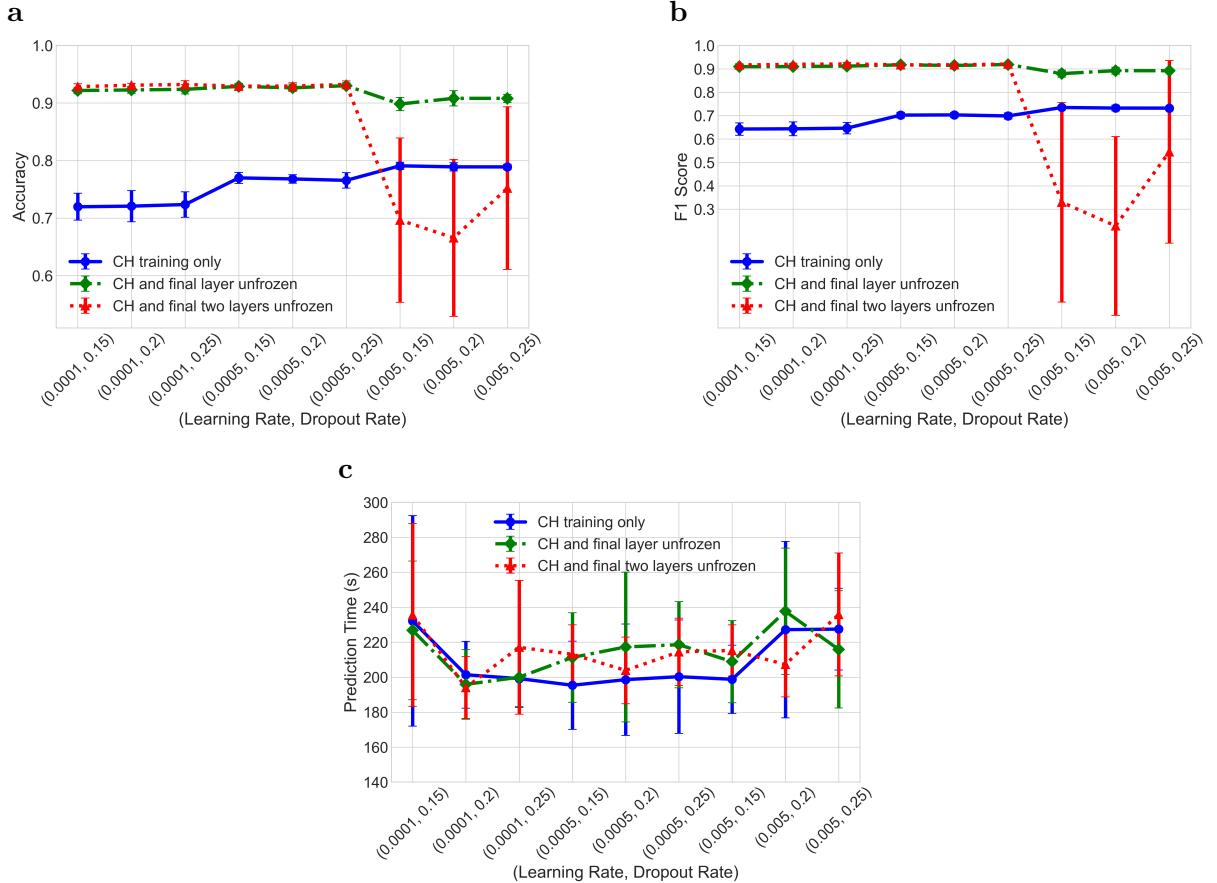


Figure (9) Results of prediction on In-house Two data with the domain adapted models optimised with AdamW optimiser: (a) Prediction accuracy; (b) F-1 score; (c) Time taken on CPU (s)

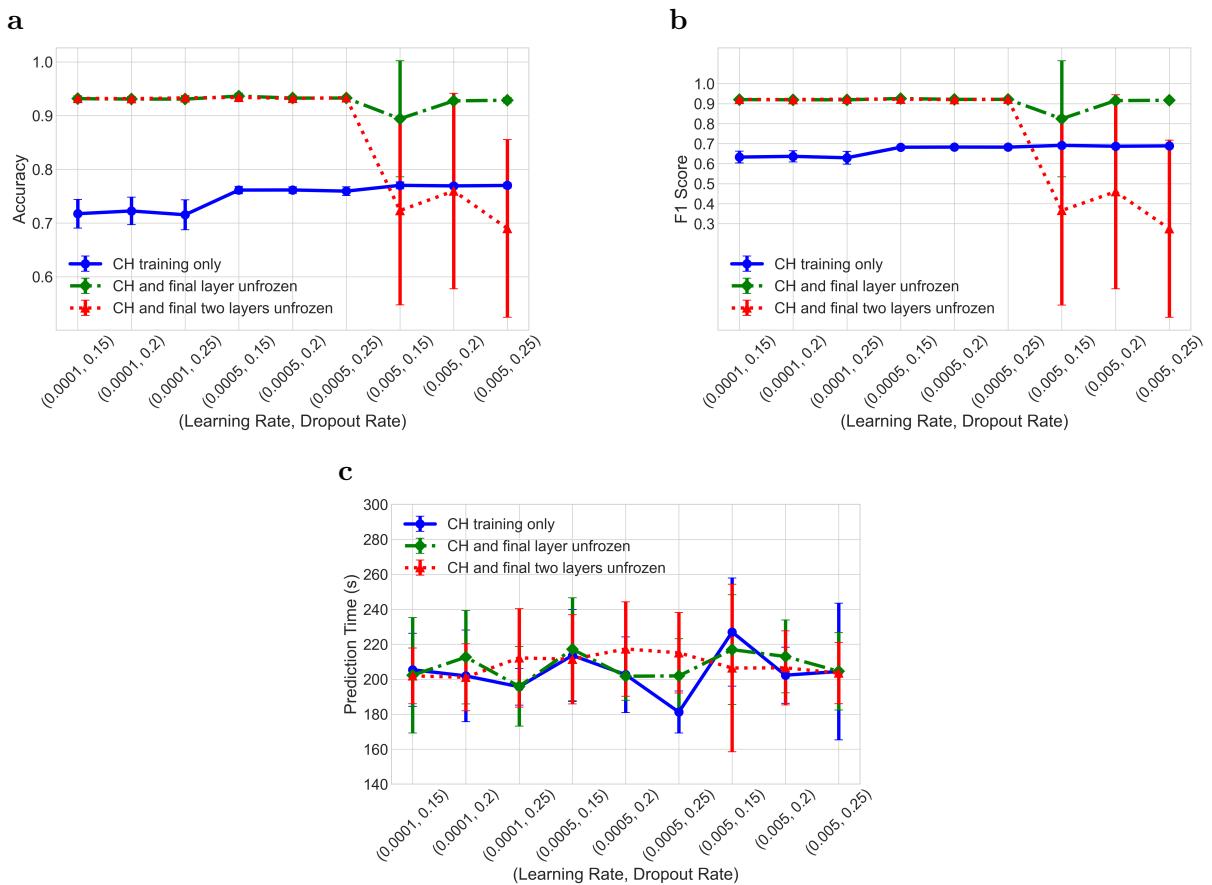


Figure (10) Results of prediction on In-house Two data with the domain adapted models optimised with Adam optimiser: (a) Prediction accuracy; (b) F-1 score; (c) Time taken on CPU (s)