NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
**College of Engineering**

Automated Essay Scoring

with Incremental Learning Through Crowdsourcing

Student: Huang Zhilin   Supervisor: Assoc Prof. Hui Siu Cheung

Motivation & Proposed Approach

Motivated by the issues of model training time overhead, complications in feature engineering, and the shortage of training data in the area of Automated Essay Scoring (AES), this research proposes a novel word-character fusion mechanism based deep learning AES model with an Incremental Learning (IL-AES) approach through crowdsourcing.
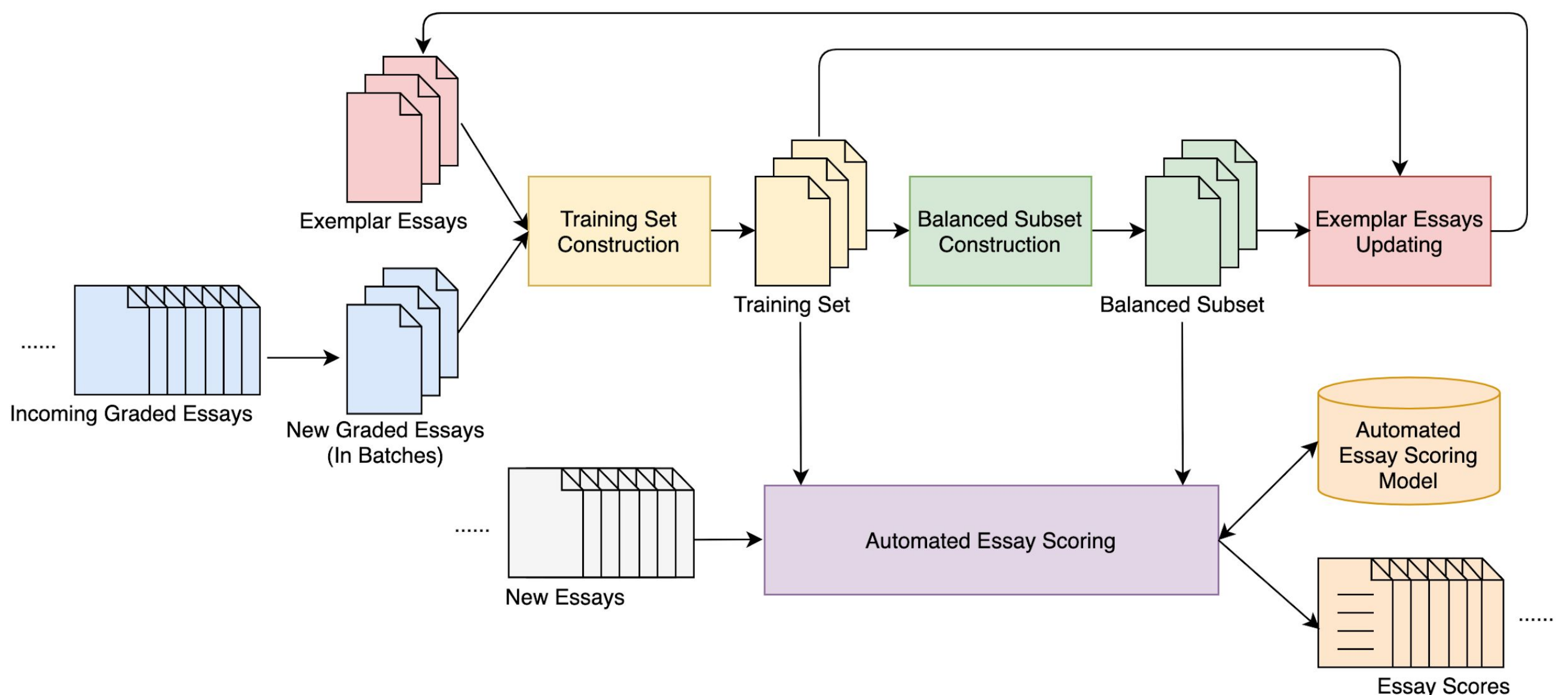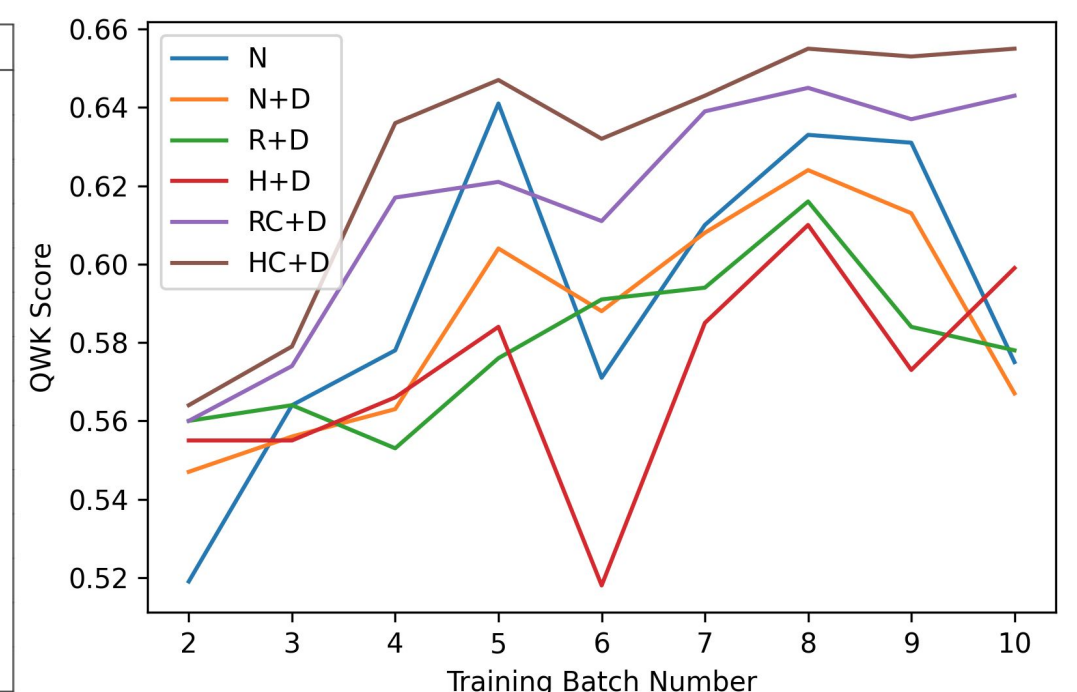

Experimental Results

Experiments were conducted using the Automated Student Assessment Prize (ASAP) dataset by Kaggle. The table below shows the experimental results (QWK scores) of the proposed AES model with word-character fusion mechanism and the baseline models.

Model	P1	P2	P3	P4	P5	P6	P7	P8	Average
Char RNN [7]	0.640	0.525	0.318	0.414	0.636	0.488	0.324	0.294	0.455
Word RNN [7]	0.680	0.589	0.463	0.592	0.657	0.545	0.544	0.419	0.561
Char GRU [7]	0.740	0.645	0.462	0.618	0.687	0.576	0.625	0.463	0.602
Word GRU [7]	0.744	0.635	0.480	0.702	0.683	0.719	0.689	0.544	0.649
Char LSTM [7]	0.741	0.648	0.442	0.612	0.711	0.574	0.621	0.470	0.602
Word LSTM [7]	0.747	0.651	0.501	0.723	0.695	0.724	0.709	0.579	0.666
Char CNN [7]	0.797	0.689	0.538	0.683	0.743	0.640	0.701	0.567	0.670
Word CNN [7]	0.805	0.690	0.598	0.763	0.732	0.780	0.792	0.688	0.731
Concat CNN [7]	0.812	**0.698**	0.594	0.765	0.750	0.773	0.792	0.686	0.734
EASE (SVR)+ [7]	0.781	0.621	0.630	0.749	0.782	0.771	0.727	0.534	0.699
EASE (BLRR)+ [7]	0.761	0.606	0.621	0.742	0.784	0.775	0.730	0.617	0.705
SKIPFLOW LSTM [9]	0.832	0.684	0.695	0.788	**0.815**	0.810	0.800	**0.697**	0.764
BERT [13]	0.792	0.680	**0.715**	0.800	0.806	0.805	0.785	0.596	0.748
BERT Ensemble [13]	0.802	0.672	0.708	0.816	0.806	**0.815**	**0.804**	0.597	0.752
WordChar	**0.852**	0.694	0.697	**0.819**	0.799	0.810	0.793	0.656	**0.765**
TSLF-ALL-Scale# [11]	0.852*	0.736*	0.731*	0.801	0.823	0.792	0.762*	0.684*	0.773*
WordChar-Scale#	0.852*	0.731	0.699	0.852*	0.832*	0.832*	0.760	0.606	0.770

Approach	2	3	4	5	6	7	8	9	10	ρ	τ
N	0.519	0.564	0.578	0.641	0.571	0.610	0.633	0.631	0.575	0.624	0.467
N+D	0.547	0.556	0.563	0.604	0.588	0.608	0.624	0.613	0.567	0.782	0.644
R+D	0.560	0.564	0.553	0.576	0.591	0.594	0.616	0.584	0.578	0.770	0.600
H+D	0.555	0.555	0.566	0.584	0.518	0.585	0.610	0.573	0.599	0.709	0.600
RC+D	0.560	0.574	0.617	0.621	0.611	0.639	0.645	0.637	0.643	**0.903**	**0.778**
HC+D	**0.564**	**0.579**	**0.636**	**0.647**	**0.632**	**0.643**	**0.655**	**0.653**	**0.655**	0.903	0.778

The experimental results (Spearman's rank correlation coefficients, Kendall rank correlation coefficients, and QWK scores) of the proposed IL-AES approach are presented in the table on the left and the figure above.