

1 Prediction of self-healing in engineered cementitious composite:
2 Machine learning comparative analysis

3 Guangwei Chen^{1,2,*}, Patrick Tang¹, Shuo Chen¹, Shangyong Wang¹, and Hongzhi Cui^{3¹}

4 ¹The University of Newcastle, Callaghan, 2308, NSW, Australia

5 ²Qiannan Normal College of Nationalities, Guizhou, China

6 ³College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China

7 **Abstract**

8 Engineered cementitious composite (ECC) is a unique material which can significantly contribute to
9 self-healing behaviour based on ongoing hydration. However, it's difficult to model and predict the self-
10 healing effect of ECC. Although different machine learning (ML) algorithms have been utilized to predict
11 several properties of concrete, the application of ML on self-healing prediction is considerably rare. This
12 paper aims to provide a comparative analysis on the performance of various machine learning mod-
13 els in predicting self-healing capability of ECC. These models include four individual methods (linear
14 regression (LR), back-propagation neural network (BPNN), classification and regression tree (CART),
15 and support vector regression(SVR)), and three ensemble methods (bagging, AdaBoost, and stacking)
16 with each of the four individual models used as the based learner. A series of experimental works on
17 self-healing performance of ECC samples was conducted and the results were used to develop the ML
18 models and compare the accuracy among the ML models. Among the individual models studied, the
19 BPNN model performed the best in terms of RMSE and R^2 . In general, all ensemble methods could
20 improve the prediction performance of individual models, however the improvement varies. Among all
21 the ML models studied including both individual and ensemble methods, the Stack LR model demon-
22 strated the best prediction results on self-healing of ECC. The results concluded that the individual and
23 ensemble methods can be used to predict the self-healing of ECC. However, selecting an appropriate
24 base learner and ensemble method is critical. To improve the performance accuracy, researchers should
25 employ different ensemble methods to compare their effectiveness with different ML models.

26 **Keywords** ECC, self-healing, machine learning, ensemble method

27 **1 Introduction**

28 According to a research project commissioned by Materials for Life (M4L), the issues associated with crack-
29 ing in concrete experienced by clients, design team members and contractors were more than any other prob-
30 lems [1]. Moreover, cracks are primarily responsible for the reduction of strength and stiffness of concrete.
31 In European countries, the annual cost spent on maintenance, refurbishment, and repair of concrete cracks
32 in prolonging the service life of infrastructure is estimated around 50% of their annual construction budget
33 [2]. It has been suggested by M4L that self-healing cementitious materials is of great potential to address the
34 problems associated with concrete cracking and reduce the maintenance costs over a structure's lifetime [1].

35 The inspiration of self-healing comes from the biomimicry concept and the healing process in living
36 nature [3]. For example, the skin of humans or animals can biologically repair itself from simple injuries.
37 In cement-based materials, the process of crack self-healing can be categorised into two major mechanisms,
38 autogenous healing and autonomous healing [4]. The former indicates the self-healing ability resulted from
39 the physical and/or chemical composition of the cementitious matrix, whereas the self-healing mechanism

40 of the latter is triggered by some biological agents, such as bacteria which are deliberately introduced into
41 the cementitious matrix.

42 Generally, the autogenous self-healing of concrete is mainly controlled by two mechanisms including
43 (1) further hydration of cement particles and/or swelling of calcium silicate hydrate; (2) calcium hydroxide
44 carbonation [5, 6]. It has been reported that the crack widths of 10 μm [7], 100 μm [8], 200 μm [9], 205 μm
45 [5] and 300 μm [10] of ECC can be self-healed completely [11].

46 Engineered cementitious composite (ECC) is a high performance fiber-reinforced cementitious com-
47 posite and its matrix design is strongly associated with the autogenous self-healing mechanism [12]. ECC
48 features high tensile ductility with a typical fiber-volume fraction of 2% [13, 14] to promote the self-healing
49 ability [4].

50 However, the intrinsic self-healing ability of ECC is complex and difficult to predict because of differ-
51 ent mineral admixture types, interactivity between different composites in the cementitious matrix and its
52 interaction with the exposed environment [15], and unpredictable crack location, orientation and width [16].
53 Previous studies have explored the influence of several factors such as limestone powders (LP) [17, 18], fly
54 ash (FA) [19, 20], hydrated lime [21], water/binder ratio [22], water permeation [23] and different curing
55 conditions (air, carbon dioxide, wet/dry and water) [24] on self-healing behaviour of ECC. However, the
56 relationship between multiple factors is unclear and non-linear, so it's difficult to predict self-healing of ECC
57 mathematically based on the available data. Moreover, mathematical models based on empirical data are
58 generally in regression forms, which cannot be used when the problem (e.g. prediction of self-healing po-
59 tential of ECC) contains too many independent variables because of less accuracy and more assumption in
60 the regression form (linear, non-linear, etc.) [25].

61 To compensate for the drawbacks of mathematical models with multiple interaction variables, machine
62 learning (ML) techniques have been used for solving many civil engineering problems with multiple vari-
63 ables. They are model-free approaches that do not rely on predefined models [26]. Many research works have
64 been conducted using ML algorithms for the prediction of various properties of concrete. Gilan et al.[27] de-
65 veloped a hybrid Support Vector Regression (SVR) - Particle Swarm Optimization (PSO) algorithm model to
66 predict the compressive strength and Rapid Chloride Penetration Test (RCPT) results of concretes containing
67 metakaolin. Yan et al. [28] predicted bond strength of glass fiber-reinforced polymer bar in concrete by Arti-
68 ficial Neural Network (ANN) with Genetic Algorithm (GA). Yaseen et al.[29] proposed a ML method called
69 Extreme Learning Machine (ELM) to predict the compressive strength of lightweight foamed concrete.

70 In the literature, the performance of various ML algorithms in predicting concrete properties have been
71 evaluated and compared. Yan and Shi [30] reported that SVR is better than other individual methods in
72 predicting elastic modulus of normal and high strength concrete. Chou [31] compared the performance of
73 individual and ensemble methods for predicting the mechanical properties of high performance concrete, the
74 results reveal that ensemble learning strategies outperform individual learning techniques in predicting high
75 performance concrete compressive strength. Reuter et al. [26] employed three individual approaches for
76 modeling concrete failure surfaces, the three approaches are able to fit the experimental data with low error.
77 Sobhani et al. [32] suggested that ANN and a proposed fuzzy inference system are more reliable than tra-
78 ditional regression models on predicting no-slump concrete. Omran et al. [33] compared the compressive
79 strengths of an environmentally friendly concrete predicted by using three individual methods, two ensemble
80 methods, and four regression tree models. Their results showed that the individual gaussian process regres-
81 sion model and its related ensemble models outperformed other models for compressive strength prediction
82 of concrete.

83 Although different ML algorithms have been utilized to predict several properties of concrete, the appli-
84 cation of ML on self-healing prediction is considerably rare. Recently, Mauludin and Oucif [34] reviewed
85 the common methods used for modeling autogenous self-healing of concrete, and stated that the methods
86 can be classified into two categories: (1) numerical simulation and (2) ML. However, the only ML model
87 reviewed in their study was the GA-ANN method proposed by Ramadan et al. [3] to predict the self-healing
88 ability of cement-based materials using a dataset collected from literature. The results showed that the

89 GA-ANN model was capable of capturing the complex effects of various self-healing agents (e.g., biochemical
90 material, silica-based additive, expansive and crystalline components) on self-healing performance of
91 cement-based materials.

92 Wang et al. [35] reported several ML models and an online ensemble learning framework to optimize
93 the predictive computational models for prediction of self-healing efficiency. The effectiveness of the online
94 ensemble learning framework was studied experimentally. More recently, Chaitanya et al. [36] used an
95 ANN model to predict the self-healing property of concrete containing ground granulated blast furnace slag
96 in terms of compressive strength recovery based on 51 samples collected from their experimental studies.
97 Generally, the predicted results obtained from the ANN model were in good agreement with the experimental
98 values. Zhuang and Zhou [37] conducted a comparative study on six ML algorithms including SVR, Decision
99 Tree Regression (DTR), Gradient Boosting Regression (GBR), ANN, Bayesian Ridge Regression (BRR) and
100 Kernel Ridge Regression (KRR) for crack-repairing capacity of the bacteria-based self-healing concrete. The
101 results showed that GBR performed much better than other models with the R^2 values of 0.93 and 0.74 for
102 the training set and testing set, respectively. However, the R^2 values of most models were less than 0.7 on
103 both training and testing sets. Although extensive experiments with different combinations of influencing
104 variables were utilized to generate the empirical dataset, their study only selected three variables including
105 the number of bacteria, the healing time and the initial crack width to predict the crack closure percentage as
106 the output.

107 To the best of our knowledge, there has been no study to date to predict the self-healing of ECC using
108 ML approach. The information about prediction performance of individual and ensemble ML models on
109 self-healing of ECC would be very useful to the design of ECC with self-healing capacity. Thus, this study
110 aims to provide a comparative analysis on the performance of various ML models in predicting self-healing
111 capability of ECC. The ML model with the best performance can be used as a baseline prediction model for
112 developing advanced models in the future.

113 In this paper, four ML individual methods including linear regression (LR), SVR, back-propagation neural
114 network (BPNN), and classification and regression tree (CART) were proposed to predict the self-healing
115 capability of ECC. To improve prediction accuracy, three ensemble methods namely bagging, AdaBoost and
116 stacking were used to construct ensemble models using the individual models as the base learners. A series of
117 experimental works on self-healing performance of ECC samples was conducted and the results were used to
118 develop the ML models and compare the accuracy among the ML models. Experimental data collected from
119 the experiments were first preprocessed and then divided into a 10-fold cross-validation algorithm (details
120 refer to Section 4.1) to avoid overfitting. Figure 1 summarizes the steps that were performed when predicting
121 the self-healing capability of ECC.

122 This paper is organized as follows. Section 2 presents the experimental program detailing the materials
123 used for ECC specimen preparation and the test set-up for crack data measurement. The concepts and
124 formulations of individual and ensemble models used for predicting the self-healing capability of ECC are
125 presented in Section 3, whereas the validation and evaluation methods are described in Section 4. In Section
126 5, the computational results are presented and compared, and the model with best prediction performance is
127 identified. Finally, Section 6 draws the major conclusions from this work and suggests some directions for
128 future research.

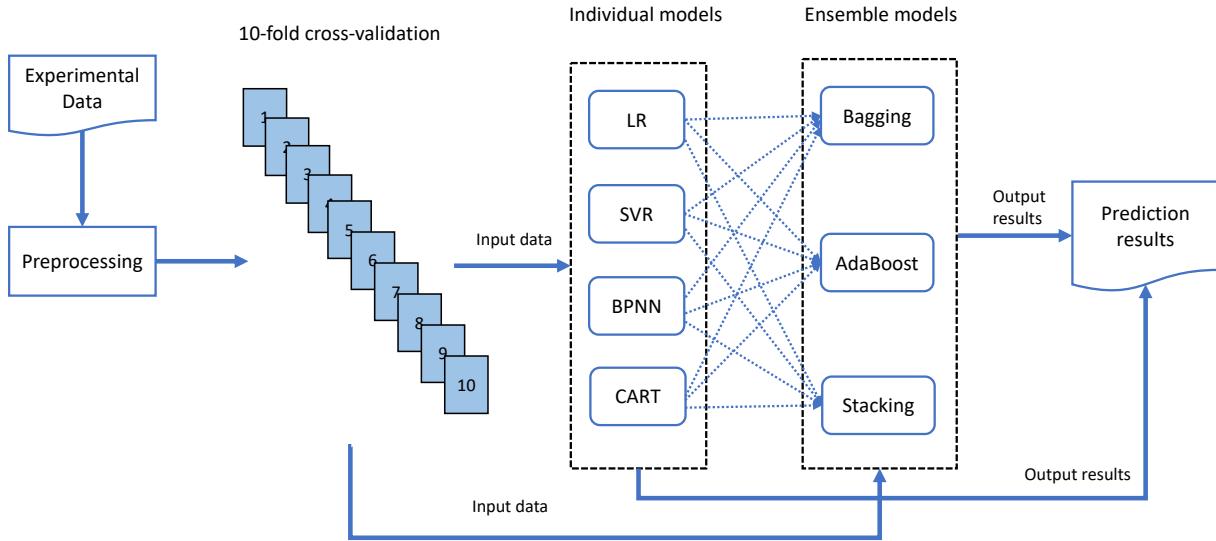


Figure 1: Flow chart of implementing prediction models for self-healing capability of ECC

129 2 Experimental Program

130 2.1 Materials and Mixture Proportion

131 In the experimental part, samples of ECC with different mineral admixtures were prepared. The materials
 132 including general purpose cement (GPC), fly ash (FA), silica fume (SF), hydrated lime powder(LP), fine
 133 sand, polyvinyl alcohol (PVA) fibers, as well as water and high range water reducing admixture (HRWR)
 134 were used. GPC and FA were supplied by Boral in accordance with Australian Standard AS 3972-2010
 135 [38], while LP was the Adelaide Brighton Hydrated Lime with a specific gravity of 2.2-2.3, and a typical
 136 fineness of 0.1% retained on a $75 \mu\text{m}$ sieve and less than 0.05% on a $250 \mu\text{m}$ sieve. The physical and chemical
 137 properties of cementitious materials are shown in Table 1. Fine sand with an average grain size of $150 \mu\text{m}$ and
 138 a fineness modulus of 2.01 was used. The PVA fibers were supplied by Domocrete and their mechanical and
 139 geometrical properties are described in Table 2.

140 All ECC mixtures were prepared with a constant water to cementitious materials (W/CM) ratio of 0.29
 141 and a constant sand to CM (PC + FA + LP+SF) ratio of 0.36. All fine aggregates were in saturated surface
 142 dried condition prior to mixing. The abbreviations for labelling specimens were adopted in such a way that
 143 the letters FA, SF and LP stand for samples with fly ash, silica fume and limestone as binder materials,
 144 respectively. The number after the letters shows the percentage of materials into the binder system. For
 145 example, the FA70 mixture is related to an ECC sample with binder containing 70% FA by weight, whereas
 146 FA60-SF10 was the mixture with 60% FA and 10% SF. A total of nine ECC mixtures were prepared and the
 147 details of mix proportion are shown in Table 3.

148 2.2 Sample preparation and crack measurement

149 A planetary-type mixer of 50 L capacity was used to produce ECC specimens. During the mixing process,
 150 the solid ingredients including cement, mineral admixtures and sand were initially placed into the mixer and
 151 dry mixed for 30 seconds. Then, the water with HRWR was added and the mixture was mixed for 2 minutes.
 152 After that, the PVA fibers were slowly added and mixing was continued until uniform distribution of fibers in
 153 the mix. After mixing, ECC pastes were cast into standard moulds with dimension of $\varnothing 100\text{mm} \times 200\text{mm}$.

Table 1: Physical and chemical properties of cementitious materials

<i>Chemical composition (%)</i>	GPC	FA	LP	SF
Silica (SiO_2)	19.8	65.90	1.8	95.10
Alumina (Al_2O_3)	5.3	24.0	0.5	0.21
Iron oxide (Fe_2O_3)	3.0	2.87	0.6	0.29
Calcium oxide (CaO)	64.2	1.59	72.0	-
Magnesia (MgO)	1.3	0.42	1.0	-
R_2O	0.6	1.93	-	-
Sulfur trioxide (SO_3)	2.7	-	-	-
Titanium oxide (TiO_2)	0.28	0.91	-	-
Manganic oxide (Mn_2O_3)	0.22	-	-	-
Zirconia (ZrO_2) + Hafnium (HfO_2)	-	-	-	3.46
Loss on ignition (%)	2.8	1.53	24.0	1.4
Density (g/cm^3)	3.08	2.43	2.25	2.26
Specific surface area (m^2/kg)	-	655	460	1.5×10^4

Table 2: Properties of PVA

Length (mm)	Length/ diameter ratio	Young's modulus (MPa)	Elongation (%)	Tensile strength (MPa)	Density (g/cm^3)
8	200	42000	7	1600	1.3

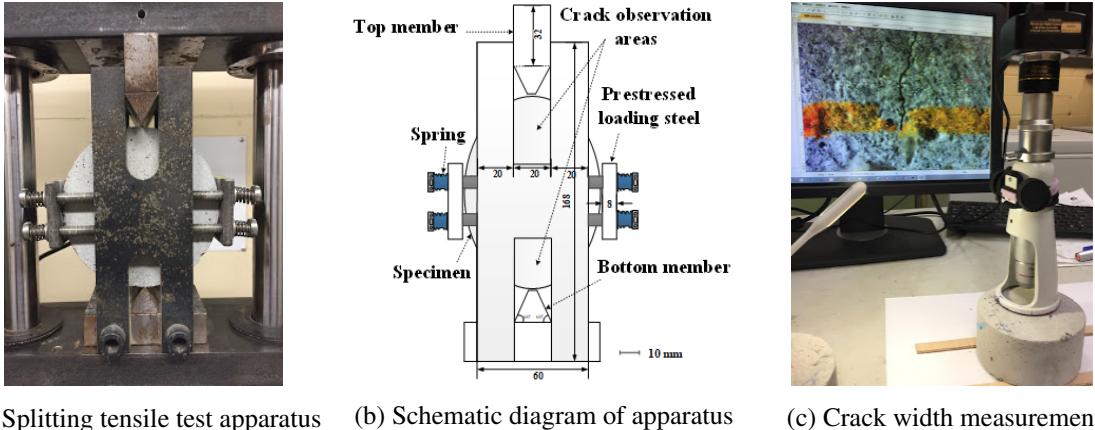
Table 3: Mix proportion of all ECC mixtures

Mix	Water/CM	Sand	Water	fibre (V)	GPC	Fly ash	SF	LP	HRWR
FA70	0.29	419.67	338.07	26	349.73	816.03	0.00	-	5.13
FA65-SF5	0.29	419.67	338.07	26	349.73	757.74	58.29	-	5.13
FA60-SF10	0.29	419.67	338.07	26	349.73	699.45	116.58	-	5.13
FA55-SF15	0.29	419.67	338.07	26	349.73	641.16	174.86	-	5.13
FA65-LP5	0.29	419.67	338.07	26	349.73	757.74	-	58.29	5.13
FA60-LP10	0.29	419.67	338.07	26	349.73	699.45	-	116.58	5.13
FA55-LP15	0.29	419.67	338.07	26	349.73	641.16	-	174.86	5.13
FA55-SF5-LP10	0.29	419.67	338.07	26	349.73	641.16	58.29	116.58	5.13
FA55-SF10-LP5	0.29	419.67	338.07	26	349.73	641.16	116.58	58.29	5.13

154 The specimens were demolded 24 hours after casting and stored in a curing room with a temperature of
 155 $23 \pm 2^\circ\text{C}$ and the relative humidity (RH) of $90 \pm 5\%$ for 28 days for 28 days. To prepare splitting tensile
 156 test samples, the cylinder specimens were cut into specimens with a diameter of 100 mm and a thickness of
 157 50 mm using a diamond blade saw.

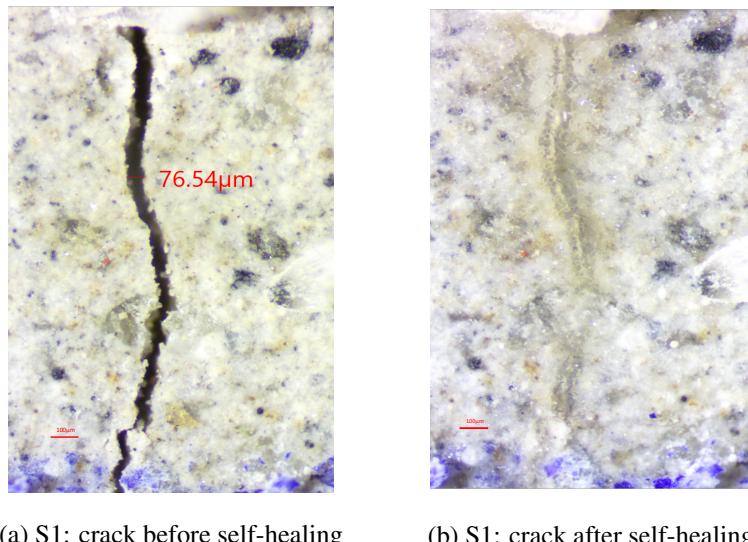
158 A newly developed splitting tensile test apparatus was used to generate micro-cracks as shown in Figure
 159 2 (a). It consisted of a steel frame, top member, bottom member, prestressed loading steel plates (5 mm
 160 thick) on both sides with loading nuts and wire springs, as shown in Figure 2 (b). Both steel plates were
 161 connected to the steel frame by nuts and wire springs. The specimen was placed inside the steel frame and
 162 then pre-stressed by the steel plates from both sides limiting the propagation and size of crack and preventing
 163 excessive crack growth.

164 Micro-cracks less than 150 μm were produced by pre-loading the ECC samples up to 70% of their
 165 maximum splitting strength. A digital microscope was used to measure the crack width on the surface of
 166 specimens as shown in Figure 2 (c). After the pre-loading, the cracked specimens were subjected to wet-dry
 167 (W/D) cycles to promote self-healing. Each W/D cycle consisted of submersion in water for 24 hours and
 168 drying in laboratory conditions at $23 \pm 2^\circ\text{C}$ and a RH of $50 \pm 5\%$ for 24 hours. After 10 W/D cycles, the
 169 cracks were measured again by the digital microscope to examine partial or full closure of crack. Figure 3
 170 illustrated the self-healing of cracks of an ECC specimen before and after the 10 W/D cycles.



(a) Splitting tensile test apparatus (b) Schematic diagram of apparatus (c) Crack width measurement

Figure 2: Splitting tensile test apparatus and microscope used in experiment for creating and measuring ECC cracks



(a) S1: crack before self-healing (b) S1: crack after self-healing

Figure 3: Comparison of crack width changes in two ECC specimens, S1 and S2, before and after self-healing

171 2.3 Data Collection

172 Experimental data for prediction were gathered with four features, including crack width before self-healing
 173 (representing the influencing factor of self-healing), and the mineral contents of FA, SF, and LP. It is note-
 174 worthy that the factors such as GPC, sand, W/CM, and healing time were kept constant and hence, they were

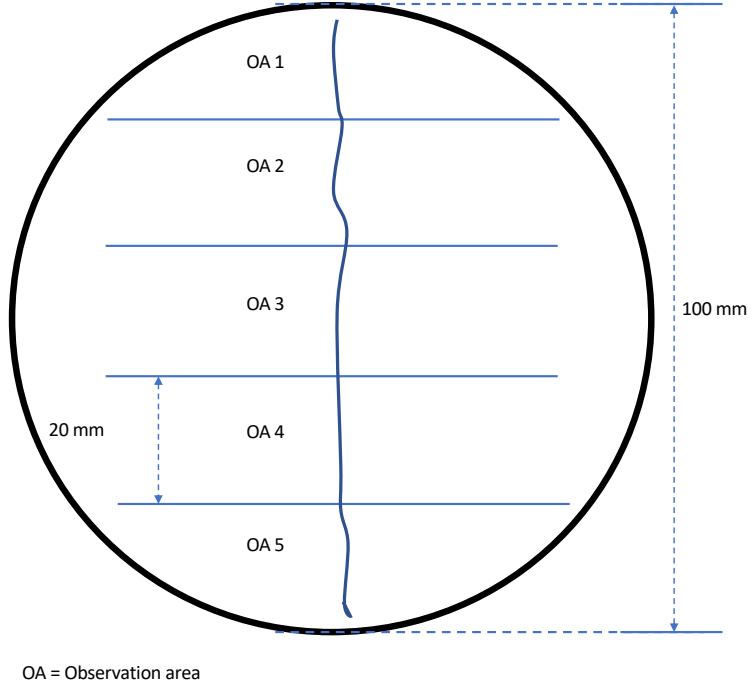


Figure 4: Schematic diagram of measuring observation areas on the surface of ECC mixture specimen

175 excluded in the prediction modeling. For each ECC mixture, there were 6 identical test specimens. After
 176 pre-loading, the crack widths of the specimens were measured using the digital microscope before and after
 177 the self-healing. Four horizontal lines were drawn on the surface of each specimen along the direction of
 178 vertical force, which divided the specimen into five observation areas. The schematic diagram of the mea-
 179 surement is shown in Figure 4. In each observation area, only one crack data was recorded if the crack width
 180 showed little or no change along the vertical force, otherwise, multiple crack data would be collected. To-
 181 tally, 617 crack data samples were collected from nine mixtures to construct the ML training-testing dataset
 182 [39]. Table 4 shows the number of collected samples and range of crack width before and after self-healing
 183 in each mixture.

184 2.4 Preprocessing of Data

185 Since the input and output data of different features vary in range and units, the features with bigger number
 186 would steer the model performance. As shown in Table 3, the range of FA varies from 641.16 to 816.03 kg,
 187 but the range of SF varies from 0 to 174.86 kg. From Table 4, the range of crack width varies from 0 to
 188 135.47 μm . To eliminate this potential bias, the experimental data was preprocessed through the min-max
 189 normalization to scale the range of all features into [0,1] with the following equation:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

190 Where x' was the scaled value of the variable x , x_{max} , x_{min} were the maximum and minimum values of
 191 variable x respectively.

Table 4: Number of crack samples and range of crack width before and after self-healing collected from the ECC mixes

Mix	Number of crack samples	Crack width before self-healing		Crack width after self-healing	
		Min (μm)	Max (μm)	Min (μm)	Max (μm)
FA70	87	3.28	134.69	0	121.37
FA65-SF5	77	4.37	135.47	0	124.01
FA60-SF10	88	5.18	121.78	0	113.11
FA55-SF15	88	3.45	115.8	0	109.53
FA65-LP5	112	7.65	119.45	0	105.65
FA60-LP10	37	5.62	126.82	0	110.97
FA55-LP15	61	6.42	132.65	0	115.95
FA55-SF5-LP10	34	8.74	123.09	0	110.78
FA55-SF10-LP5	33	4.64	131.57	0	119.79

192 3 Proposed Machine Learning Models

193 To predict the self-healing capability of ECC, four individual ML models including LR, SVR, BPNN and
 194 CART, and three ensemble methods including bagging, AdaBoost and stacking were proposed. Ensemble
 195 models were constructed using individual models as the base estimators. To establish a baseline for compar-
 196 ison, the modeling parameters were set to be the same in both individual models and ensemble models. The
 197 reason for choosing these techniques was due to their popularity and some of them were even recognized
 198 as the top data mining algorithms in related fields of concrete [31]. The proposed individual and ensemble
 199 techniques are described in the following subsections.

200 3.1 Linear Regression

201 LR attempts to determine the relationship between a dependent variable (response variable) and one or more
 202 independent variables (explanatory variables) by fitting a linear regression equation [40]. Given our dataset
 203 $T = \{(x_i, y_i), i = 1, 2, \dots, n\}$, where $n = 617$ was the size of sample dataset. $x_i \in R^n$ was independent
 204 variables representing a sample of selected features from FA, SF, LP and crack width before self-healing, R^n
 205 was n -dimensional space, $y_i \in R^1$ was the target output (crack width after self-healing) that corresponded
 206 to x_i . Let $d = 4$ denote the number of an independent variable of a random vector $x = \{x_1; x_2; \dots; x_d\}$, and
 207 y was the corresponding output (dependent variable). The general formula of LR for predicting self-healing
 208 capability of ECC can be expressed as follows:

$$y = w_1 x_1 + w_2 x_2 + \dots + w_d x_d + b \quad (2)$$

209 where $w_i, (i = 1, 2, \dots, d)$ was denoted as the regression coefficient, b was an error term. The prediction
 210 performance of LR was used as a benchmark to compare the performance of other individual and ensemble
 211 models in this study.

212 3.2 Support Vector Regression

213 The support vector machine (SVM) is a supervised machine learning method first introduced by Vapnik [41,
 214 42] based on statistical learning theory [43]. Since then, it has gained popularity due to attractive features,
 215 and promising empirical performance. SVM includes two main categories: support vector classification

216 (SVC) and SVR. For classification purposes, SVMs often used a *kernel* function to map the input data as
 217 vectors to a high-dimensional feature space so that an optimal separating hyperplane can be constructed [44].

218 For regression purposes, the basic idea is to provide a nonlinear function by mapping input data into a
 219 high-dimensional feature space, where a special type of hyperplane is constructed. After that, a regression
 220 model is established in the hyperplane [45].

221 Given our dataset $T = \{(x_i, y_i), i = 1, 2, \dots, n\}$, where $n = 617$ was the size of sample dataset, $x_i \in R^n$
 222 was the input vector representing selected features of a sample, including FA, SF, LP and crack width before
 223 self-healing, R^n was the n -dimensional vector space, $y_i \in R^1$ was the target output indicating crack width
 224 after self-healing that corresponded to x_i . The SVR aimed to seek an optimum regression function $f(x)$ with
 225 minimal empirical risk, which can be expressed as follow:

$$f(x) = \langle w, x \rangle + b \quad \text{with } w \in T, b \in R \quad (3)$$

226 where $\langle \cdot, \cdot \rangle$ was denoted as the dot product in T , w and b were the weight vector and bias value which are
 227 estimated by minimizing the empirical risk, that was, the distance between the predicted crack width and the
 228 target crack width after self-healing.

229 SVR adopts an ϵ -insensitive loss function penalizing predictions that has a distance between the predicted
 230 crack width and the target crack width when the self-healing is greater than ϵ . Therefore, the problem of
 231 finding w and b to reduce the empirical risk with respect to an ϵ -insensitive loss function is equivalent to the
 232 convex optimization problem that minimizes the margin (w) with the full prediction error within the range
 233 of ϵ . Then this problem can be expressed as:

$$\begin{aligned} & \text{minimize} \quad \frac{1}{2} \|w\|^2 \\ & \text{subject to} \quad \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon \\ \langle w, x_i \rangle + b - y_i \leq \epsilon \end{cases} \end{aligned} \quad (4)$$

234 By introducing slack variables ξ, ξ_i^* to allow some errors to cope with infeasible solution of the optimization
 235 problem, the formulation can be generated as [42]:

$$\begin{aligned} & \text{minimize} \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ & \text{subject to} \quad \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \quad (5)$$

236 The constant C was the penalty value imposed on predictions that lied outside the ϵ margin. Lagrange
 237 multipliers are included to solve this problem. By constructing the objective function and all constraints, a
 238 dual set of variables are introduced as follows: [46]:

$$\begin{aligned} L_P = & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) - \sum_{i=1}^n (\eta_i \xi_i + \eta_i^* \xi_i^*) \\ & - \sum_{i=1}^n \alpha_i (\epsilon + \xi_i - y_i + \langle w, x_i \rangle + b) \\ & - \sum_{i=1}^n \alpha_i^* (\epsilon + \xi_i^* + y_i - \langle w, x_i \rangle - b) \\ s.t. \quad & \alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0 \end{aligned} \quad (6)$$

239 Where L_P was the Lagrangian and $\alpha_i, \alpha_i^*, \eta_i, \eta_i^*$ were Lagrange multipliers.

240 The optimality can be achieved by the partial derivatives of L_P with respect to the primal variables
241 following the saddle point condition. Then the function of SVR is obtained as:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \quad (7)$$

242 As for the nonlinear regression, the input data have to be mapped into a high-dimensional feature space,
243 in which the dot product can be replaced by a kernel function $k(x_i, x_j) = \phi(x_i)^T \phi(x_j)$, and the function (7)
244 can be written as:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (8)$$

245 Different SVM algorithms use differing kinds of kernel functions such as linear, polynomial, radial basis
246 function and sigmoid kernel. In this work, the Gaussian radial basis function (RBF) was chosen, which was
247 defined as [47]:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (9)$$

248 3.3 Artificial Neural Network

249 Artificial neural network (ANN), also called neural network, is originated from simulating biological neural
250 networks. Generally, it consists of many neurons in layers including one input layer, one or several hidden
251 layers and an output layer [48]. The neurons are fully interconnected between the neighboring layers by
252 weight, and typically no inter-connections between neurons within the same layer [49].

253 There are many possible network structures available, BPNN was utilized in this study because of back-
254 propagation (BP) algorithms is the most widely used and effective learning algorithm for training an ANN.
255 A preliminary architecture of the BPNN was determined to be 4 - n - 1, where 4 input neurons represented
256 the input features standing for FA, LP, SF and crack width before self-healing, n = 5 indicated the num-
257 ber of neurons in the hidden layer, and 1 target neuron in the output layer for the predicted crack width
258 after self-healing. This is a three-layer network with one hidden layer capable to approximate most contin-
259 uous functions, of which the complex nonlinear relationship could be approximated in accuracy [28]. The
260 architecture of the BPNN model for predicting self-healing is demonstrated in Figure 5.

261 Given a set of inputs $\{x_1, x_2, x_3, \dots, x_n\}$, while information was passed through the input layer to the
262 hidden layer, each neuron in the input layer was multiplied by respective weights added by a bias and are
263 summed together. After that, an activation function f was applied to form the output z . This can be expressed
264 in the following equation [25]:

$$z = f\left(\sum_{i=1}^n w_{ij} x_i + b_j\right) \quad (10)$$

265 where w_{ij} was the connection weights between the i th neuron of input and the j th neuron in the hidden
266 layer, and b_j was the bias of the j th neuron. The sigmoid function was applied as the activation function
267 between the input, hidden, and output neurons to form the output.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (11)$$

268 The goal of training a neural network is to determine the values of the connection weights and the biases
269 of the neurons. The back propagation indicates an iterated method to adjust the weights from output layer

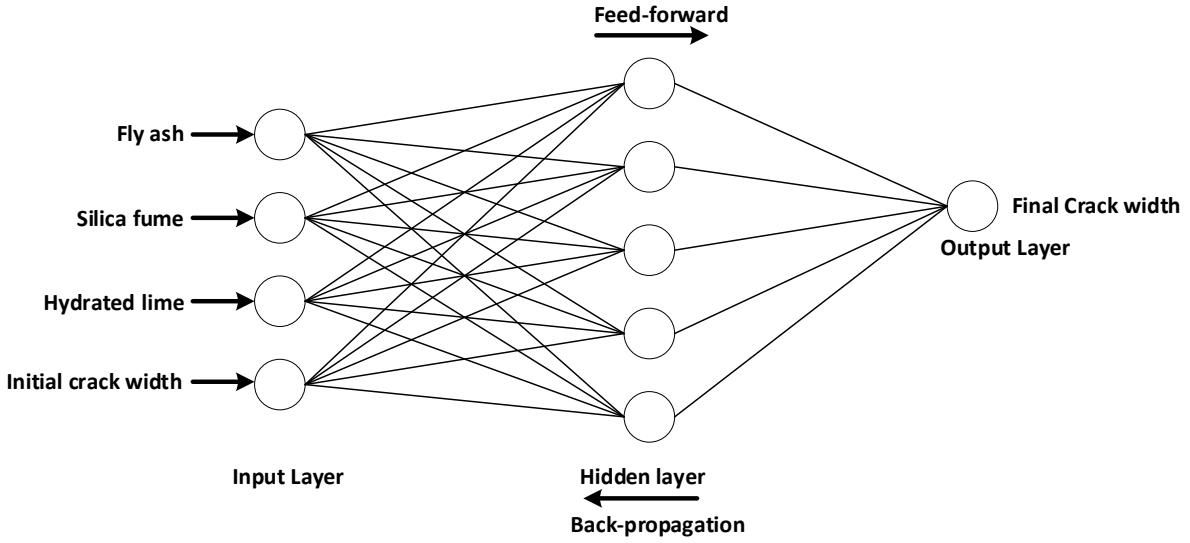


Figure 5: Schematic diagram of BPNN model for predicting self-healing capability of ECC

270 to input layer. At first, the outputs were calculated feed-forward from the input layer via the hidden layer
 271 to the output layer. Then an error was generated by comparing the output with the target output. After that,
 272 the error was back propagated to the hidden layer and input layer. By adjusting the connection weights and
 273 biases, the error was further reduced. The process was repeated until the error was minimised or reaching
 274 the termination to avoid over-fitting.

275 3.4 Classification and Regression Tree

276 The CART [50] is a tree decision algorithm that splits data into mutually exclusive subgroups based on
 277 recursive binary partitioning procedure. It develops the relationship between the target variables (the crack
 278 width after self-healing of ECC) and the independent variables (the input features of FA, SF, LP and crack
 279 width before self-healing of ECC) to create decision rules to form subgroups as branches and leaves as shown
 280 in Figure 6. The process of CART starts from the root node which contains the entire data set to construct
 281 two sub-nodes representing two categories. Then this recursion process is applied to each sub-node until all
 282 divided sub-nodes are leaf nodes. The CART tree can be either a classification tree [51] or regression tree
 283 [52] depending on the type of target and independent variables which may be categorical or numerical.

284 The key idea of constructing a CART tree is achieved by selecting a variable at each node that best splits
 285 the empirical data. To locate splits, *Gini* index was used to measure the impurity of the two child nodes
 286 containing subsets of data that were as homogeneous as possible with respect to the target variable.

287 Given a dataset had K classes and the probability of a record in the dataset which belongs to class i is
 288 $p_i, i \in \{1, 2, 3, \dots, K\}$, the *Gini* impurity can be expressed as:

$$G(p) = \sum_{i=1}^K p_i(1 - p_i) = 1 - \sum_{i=1}^K p_i^2 \quad (12)$$

289 3.5 Ensemble Methods

290 In contrast to many ML approaches such as SVM and CART (which develop a single learner from training
 291 data), ensemble methods train multiple base learners and combine them [31] to improve generalizability

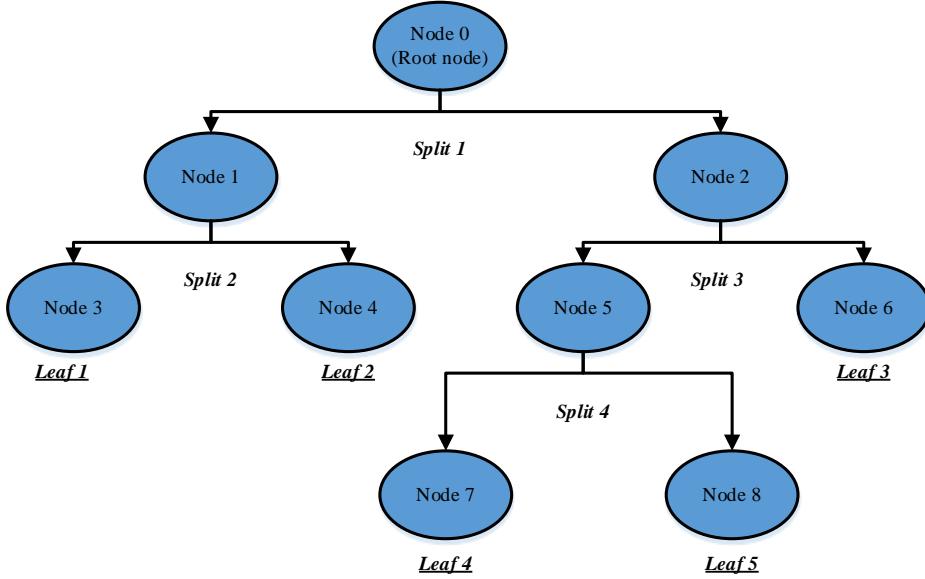


Figure 6: Structure of a classification and regression tree [52]

292 over a single estimator. Therefore, weak learners (base learners) can be boosted to become strong learner
 293 [53] in an ensemble method. The base learners in an ensemble were developed from an individual learning
 294 algorithm such as decision tree, SVM, or other kinds of learning algorithms. Breiman [54] showed that
 295 ensemble methods are usually more accurate than individual learning methods.

296 The input features of FA, SF, LP, and crack width before self-healing of ECC were considered as the
 297 d -dimensional predictor variable X , whereas, the crack widths after self-healing of ECC were the one dimen-
 298 sional output Y . Each estimator used an individual algorithm to provide one estimated function $g(\cdot)$.
 299 The output presented by ensemble-based function $g_{\text{en}}(\cdot)$ was obtained by a linear combination of individual
 300 functions. This ensemble approach can be expressed mathematically as:

$$g_{\text{en}}(\cdot) = \sum_{j=1}^N c_j * g_j(\cdot) \quad (13)$$

301 Where c_j expressed as the combination coefficients, dependent on the used ensemble models.

302 3.5.1 Bagging

303 Bagging method (bootstrap aggregating) can generate multiple versions of a predictor to obtain an aggre-
 304 gated predictor [55]. It generates multiple models independently on different versions of dataset via random
 305 bootstrapping of the original training set. In other words, several training examples could repeatedly ap-
 306 pear in different bootstrap replicates. Then the individual predictions are aggregated through a combination
 307 method (either voting or averaging) to form the final prediction. Bagging method can be used to reduce the
 308 variance of a base estimator (e.g. a regression tree), by introducing randomization into its construction pro-
 309 cedure and making an ensemble out of it. This study used four individual models to build bagging ensemble
 310 models including a LR bagging ensemble model (abbreviated as Bag_LR), a SVR bagging ensemble model
 311 (abbreviated as Bag_SVR), a BPNN bagging ensemble model (abbreviated as Bag_BPNN), and a CART
 312 bagging ensemble model (abbreviated as Bag_CART).

313 **3.5.2 AdaBoost**

314 Similar to bagging, AdaBoost method [56] manipulates the training examples to generate multiple predictions
315 to form the final prediction. The main difference with bagging is that AdaBoost applies a weight to each
316 of the training examples. In each iteration, the weights are individually updated to minimize the weighted
317 error on the training set. For example, weights on those training examples incorrectly predicted in previous
318 iteration increase, whereas the weights of the correctly predicted training examples decrease. Therefore,
319 AdaBoost tends to construct progressively more difficult learning problems in subsequent iterations. Once
320 the training process has finished, the predictions are combined through a weighted majority vote (or sum) to
321 produce the final prediction. So, the final classifier usually can achieve a high degree of accuracy in the test
322 set.

323 By combining four individual models as base estimators in AdaBoost, this study obtained four AdaBoost
324 ensemble models. They are a LR AdaBoost ensemble model (abbreviated as Ada_LR), a SVR AdaBoost en-
325 semble model (abbreviated as Ada_SVR), a BPNN AdaBoost ensemble model (abbreviated as Ada_BPNN),
326 and a CART AdaBoost ensemble model (abbreviated as Ada_CART).

327 **3.5.3 Stacking**

328 Stacking regression combines multiple regression models via a meta-regressor, using out-of-fold prediction
329 concept [57]. The stacking method used in this work splits the data set into k folds, in which the k-1 folds are
330 used to train the first level regressors in k successive rounds. In each round, the first level regressors are used
331 to predict based on the remaining 1 subset. After that, the prediction results are used and stacked as input
332 data to the second level regressors to form a final set of predictions [58]. The schematic diagram of stacking
333 model is shown in Figure 7. In this study, one stacking based ensemble model (abbreviated as Stack_LR)
334 was proposed based on two levels scheme. SVR, BPNN and CART were used as regression models in the
335 first level to get the prediction results, and LR was used as meta-regressor in the second level to combine and
336 generate the final prediction results.

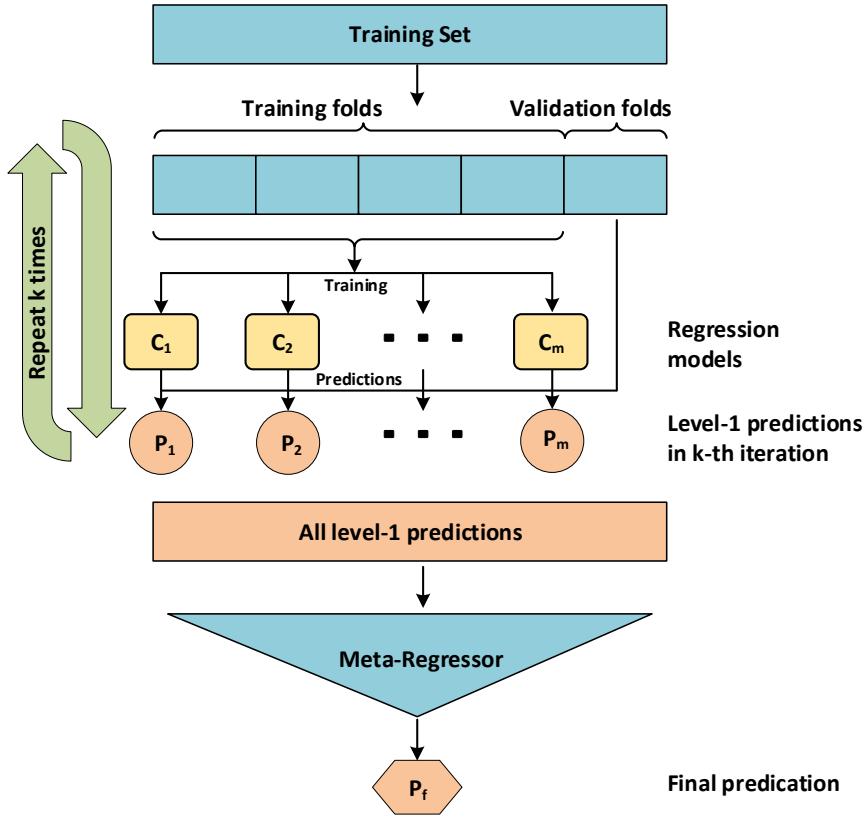


Figure 7: Schematic diagram of Stacking model [58]

337 4 Validation and Evaluation

338 4.1 Cross-validation Method

339 Generally, dataset is split to generate a training subset and a validation subset keeping the properties of the
 340 original dataset as much as possible to avoid misleading estimates. To minimize bias of random data splitting,
 341 the K-fold cross-validation is commonly used as it can yield optimal computational time and reliable variance
 342 [31, 59]. In this study, a ten-fold cross-validation approach was applied to assess model performance as
 343 shown in Figure 8. The dataset was split randomly into 10 equal-size subsets with a similar distribution. In
 344 each validation process, nine of the subsets were used for training and the rest for testing. The process was
 345 repeated 10 times [60]. The average accuracy after 10 times validation was reported as the model accuracy.

346 4.2 Performance Evaluation

347 To show and validate the accuracy of the proposed ML models, three statistical indices namely mean ab-
 348 solute Error (MAE), root mean square error (RMSE), and the coefficient of determination R^2 were used
 349 and expressed in equations (14), (15), and (16), respectively. The average deviation of the performance of
 350 an individual model or an ensemble model from a benchmark model in terms of three statistical measures
 351 (MAE, RMSE and R^2) was calculated using equation (17).

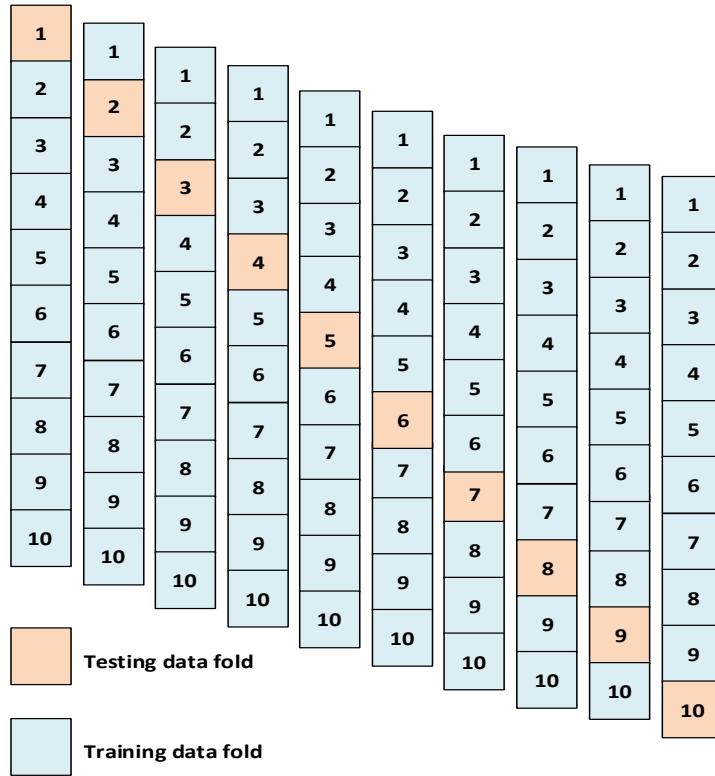


Figure 8: Ten-fold cross-validation approach

352

- Mean absolute error (MAE).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \quad (14)$$

353

- Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \quad (15)$$

354

- Coefficient of determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (16)$$

355

- Deviation (Dev)

$$Dev(\%) = \frac{P_i - P_j}{P_j} * 100 \quad (17)$$

356

Where y_i was the target output, y'_i was the predicted output, n was the number of samples, \bar{y} was the mean of the target output. Dev indicated the statistical performance improvement compared with a benchmark model, P_i was the statistical performance (MAE, RMSE or R^2) of an individual or ensemble method, and P_j

359 was the corresponding performance of a benchmark model, LR or an individual method used in the ensemble
 360 method as the base learner.

361 MAE statistics is a measure of errors between the predicted values (the estimated value of crack width of
 362 ECC after self-healing) with the target values (the observed value of crack width of ECC after self-healing in
 363 empirical data). RMSE statistics computes the square root of the average residual error between the predicted
 364 values and the target values. A lower value of MAE or RMSE indicates a better prediction performance of
 365 the model. R^2 measures the strength of association between the predicted values and the target values, based
 366 on the proportion of total variation of outcomes. A greater value close to 1 represents a better prediction
 367 performance that commendably replicates the observed crack width of ECC after self-healing. Deviation
 368 statistics indicates the improvement of the prediction performance of an individual or an ensemble model
 369 from a benchmark model that can be the LR model or the individual model used as base learners in the
 370 corresponding ensemble model.

371 5 Results and Discussion

372 In this section, the prediction performance of individual and ensemble methods are examined by MAE,
 373 RMSE and R^2 according to ten-fold cross-validation. The abbreviation for labelling models were adopted in
 374 a such a way that the letters Bag, Ada and Stack stand for the ensemble methods of Bagging, AdaBoost and
 375 Stacking, respectively. The letters LR, SVR, BPNN and CART stand for the base estimators. However, the
 376 Stack_LR model refers to combining the base methods including SVR, BPNN, and CART in the first level
 377 and using LR as a meta-regressor in the second level.

378 5.1 Prediction performance of the proposed models

379 Table 5 shows the average performance of individual and ensemble models. The ten-fold cross-validation
 380 results (MAE, RMSE, and R^2) for both individual and ensemble models and their deviation with respect to
 the results of LR model.

Table 5: Average performances of machine learning models for self-healing prediction of ECC

	Models	MAE	Dev(%)	RMSE	Dev(%)	R^2	Dev(%)
Individual models	LR	5.012	-	7.680	-	0.860	-
	BPNN	4.329	-13.6	6.515	-15.2	0.899	4.5
	CART	4.305	-14.1	6.811	-11.3	0.887	3.1
	SVR	4.296	-14.3	6.826	-11.1	0.883	2.7
Ensemble models	Ada_LR	4.784	-4.6	7.400	-3.6	0.867	0.8
	Ada_BPNN	4.226	-15.7	6.435	-16.2	0.900	4.7
	Ada_CART	4.207	-16.1	6.455	-15.9	0.898	4.4
	Ada_SVR	4.145	-17.3	6.577	-14.4	0.893	3.8
	Bag_LR	5.014	0.0	7.689	0.1	0.860	0.0
	Bag_BPNN	4.143	-17.3	6.341	-17.4	0.901	4.8
	Bag_CART	4.093	-18.3	6.358	-17.2	0.901	4.8
	Bag_SVR	4.302	-14.2	6.820	-11.2	0.883	2.7
	Stack_LR	3.934	-21.5	6.118	-20.3	0.904	5.1

381 Generally, most of the proposed models were able to learn and predict empirical data with an acceptable
 382 degree of precision. Based on the results, the Stack_LR model showed the best prediction performance
 383 as it has the highest R^2 value and lowest MAE and RMSE values. Among the individual models, SVR
 384 performed the best in terms of MAE (4.296), but BPNN has the lowest RMSE value (6.515) and highest R^2
 385 of 0.899. For the single learning based ensemble methods, Bag_CART gave the best performance in terms
 386 of 0.883 and 2.7% deviation.

387 of MAE (4.093), and Bag_BPNN performed better on RMSE value (6.341). In terms of R^2 , Bag_CART and
 388 Bag_BPNN models showed the same performance (0.901) and better than other ensemble methods except
 389 Stack_LR. The performances of all ML models described in Table 5 are depicted in Figure 9 (a), (b), and (c)
 390 in terms of MAE, RMSE and R^2 , respectively.

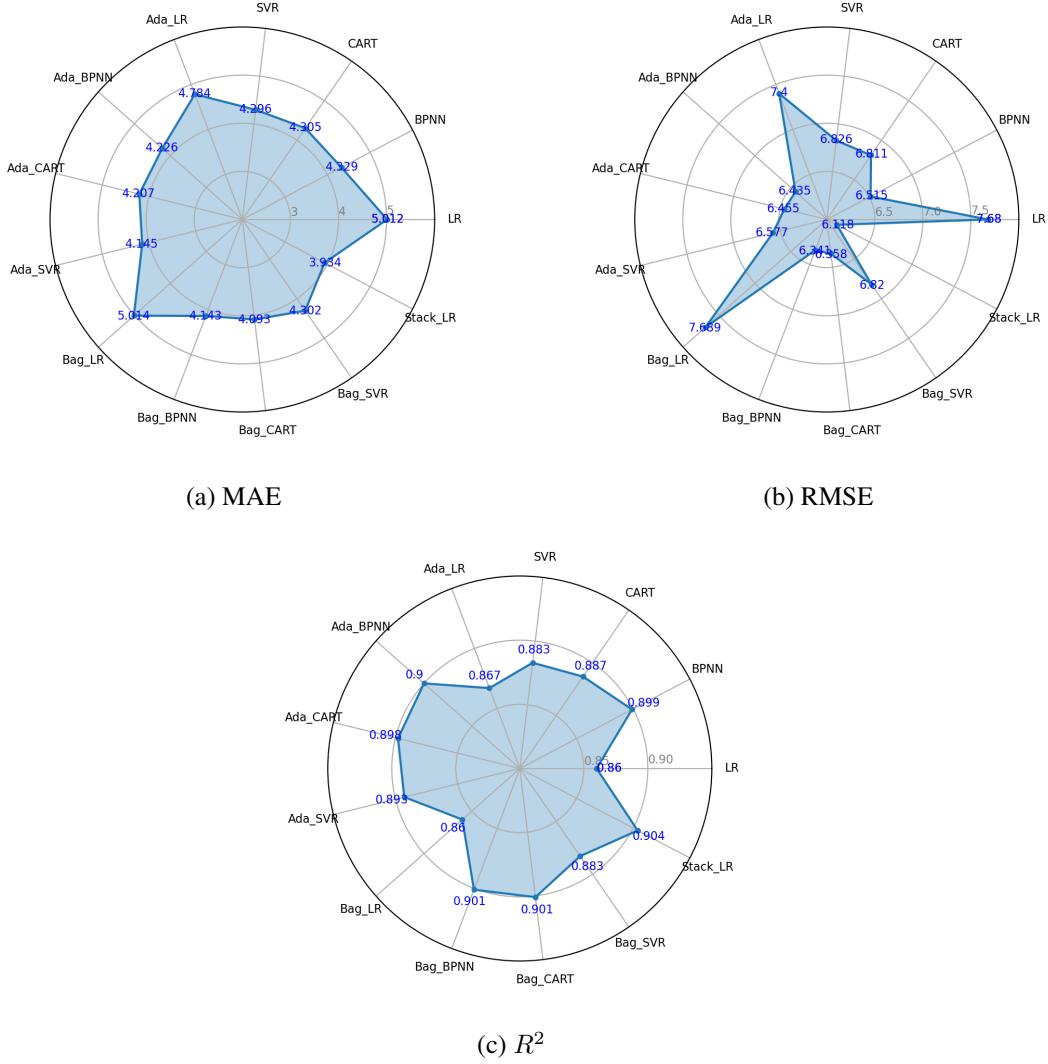


Figure 9: Average prediction performance of 10-fold cross-validation on all machine learning models for predicting self-healing ability of ECC

391 Overall, all models can noticeably reduce the error values and increase the prediction accuracy compared
 392 with LR, except Bag_LR. Among the models boosted by AdaBoost, Ada_SVR performed the best with the
 393 lowest MAE value, whereas Ada_BPNN performed the best on RMSE value showing the highest R^2 value.
 394 In case of bagging, both Bag_CART and Bag_BPNN performed better in terms MAE, RMSE and R^2 than
 395 those of the corresponding models boosted by AdaBoost. However, Bag_LR showed a poor performance
 396 compared to LR on the MAE and RMSE values. For a better comparison among the ensemble methods used,
 397 the performance results between the ensemble models and their corresponding individual (or benchmark)
 398 models are indicated in Table 6. The results indicate that most ensemble methods improved the performance
 399 of individual models. For example, the MAE and RMSE values of BPNN after bagging reduced by 4.3% and
 400 2.7%, respectively, and with a higher value of R^2 compared to those of the individual BPNN model. Among

401 all the ensemble methods studied, stacking showed the best improvement on all performance measures.

Table 6: Performance deviation of ensemble models from benchmark models on self-healing of ECC

Benchmark	Model	MAE	RMSE	R^2	Benchmark	Model	MAE	RMSE	R^2
		<i>Dev(%)</i>					<i>Dev(%)</i>		
LR	Ada_LR	-4.6	-3.6	0.8	LR	Bag_LR	0.0	0.1	0.0
BPNN	Ada_BPNN	-2.4	-1.2	0.1	BPNN	Bag_BPNN	-4.3	-2.7	0.2
CART	Ada_CART	-2.3	-5.2	1.2	CART	Bag_CART	-4.9	-6.6	1.6
SVR	Ada_SVR	-3.5	-3.6	1.1	SVR	Bag_SVR	0.1	-0.1	0.0
Ada_LR	Stack_LR	-17.8	-17.3	4.3	Bag_LR	Stack_LR	-21.5	-20.4	5.1

402 However, the results showed that the effectiveness of ensemble methods on individual models varied.
403 For instance, bagging method enhanced the performance of BPNN and CART substantially, but not for both
404 LR and SVR models. On the other hand, the AdaBoost method brought a considerable improvement for
405 LR and SVR models. To improve the performance accuracy, researchers should employ different ensemble
406 methods to compare their effectiveness on different ML models.

407 5.2 Prediction performance comparison

408 To reveal the accuracy of the proposed ML models in self-healing prediction, the variations in self-healing
409 between the observed crack widths and the crack widths predicted by the proposed ML models are shown
410 in Figures 10 to 13. Figure 10 a compares the observed results and the predicted results by using different
411 individual ML models. Figures 10 b-e show the variations in self-healing between the observed and the
412 predicted crack widths by each individual ML model corresponding to their initial crack widths before self-
413 healing. In other words, the prediction performance of models in a particular range of crack widths can be
414 revealed. It should be noted that the horizontal line located at the vertical coordinate of zero ($y = 0$) is
415 considered as the target line [25, 28]. Generally, the smaller the variation (i.e. closer to the target line), the
416 better the self-healing prediction, which means the smaller or even no variation between the observed and
417 the predicted crack widths after self-healing.

418 It has been reported that crack width control is critical to the generation and accumulation of self-healing
419 substances [5, 8, 61, 62, 63] that further affect the self-healing capability of ECC. In order to eliminate the
420 influence of cracking pattern, a newly developed splitting tensile test apparatus (shown in Figure 2) was used
421 to control crack width, length and depth, as well as the branched and accumulated crack. As shown in Table
422 4, and Figures 10 to 13, the maximum crack width before self-healing was $135.47 \mu\text{m}$ (less than $150 \mu\text{m}$),
423 and the majority of cracks were within $100 \mu\text{m}$, with most of them were less than $60 \mu\text{m}$.

424 As shown in Figure 10, the SVR model generally exhibited better prediction results than other individual
425 models, while the LR model is the worst showing substantial deviation from the target line (denoting relative
426 large differences between the observed and the predicted crack widths). For the initial crack widths less than
427 $20 \mu\text{m}$ and over $100 \mu\text{m}$ before self-healing, the variations shown in SVR model that are smaller than those
428 presented in other models had the average absolute error of 1.357 and 2.724. However, the CART model
429 showed smaller variation with average absolute error of 5.045 for the initial crack widths between 20 and
430 $60 \mu\text{m}$, while BPNN model showed the smaller variation with average absolute error of 9.565 for the initial
431 crack widths between 60 and $100 \mu\text{m}$ before self-healing. In terms of accuracy among the individual models,
432 SVR performed the best, followed by CART, BPNN and LR. This is consistent with the results shown in
433 Table 5, the MAE results of whole dataset indicated the SVR had the lowest value of 4.296 followed by
434 CART with 4.305, BPNN with 4.329 and LR with 5.012.

435 The performance of ensemble methods using Bagging and Adaboost are shown in Figures 11 and 12. In
436 general, the ensemble models bag_CART and Ada_CART exhibited lower variations in self healing results

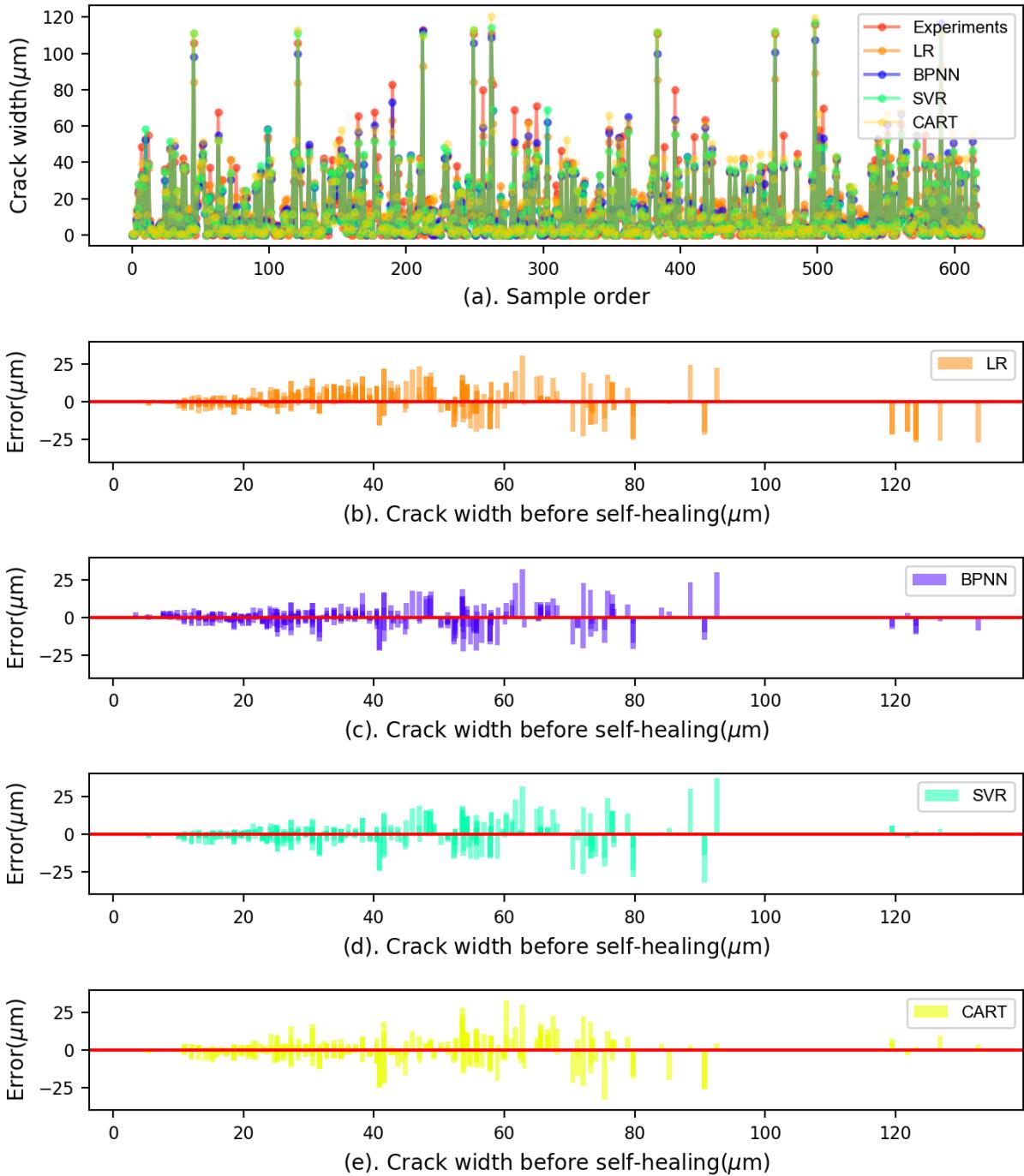


Figure 10: Comparison of predicted and observed crack width after self-healing of ECC for individual models

437 compared to other ensemble models. In particular, the average absolute error of 5.000 and 5.037 observed
 438 for bag_CART and Ada_CART for crack widths between 20 and 60 μm are smaller than CART with 5.045
 439 shown in Figure 10 e. This is also observed in the MAE values (between the predicted crack width and
 440 the experimental observed crack width) of whole dataset that are 4.093, 4.207 and 4.305 for bag_CART,
 441 Ada_CART and CART, respectively. However, the variations among BPNN, Ada_BPNN and bag_BPNN are

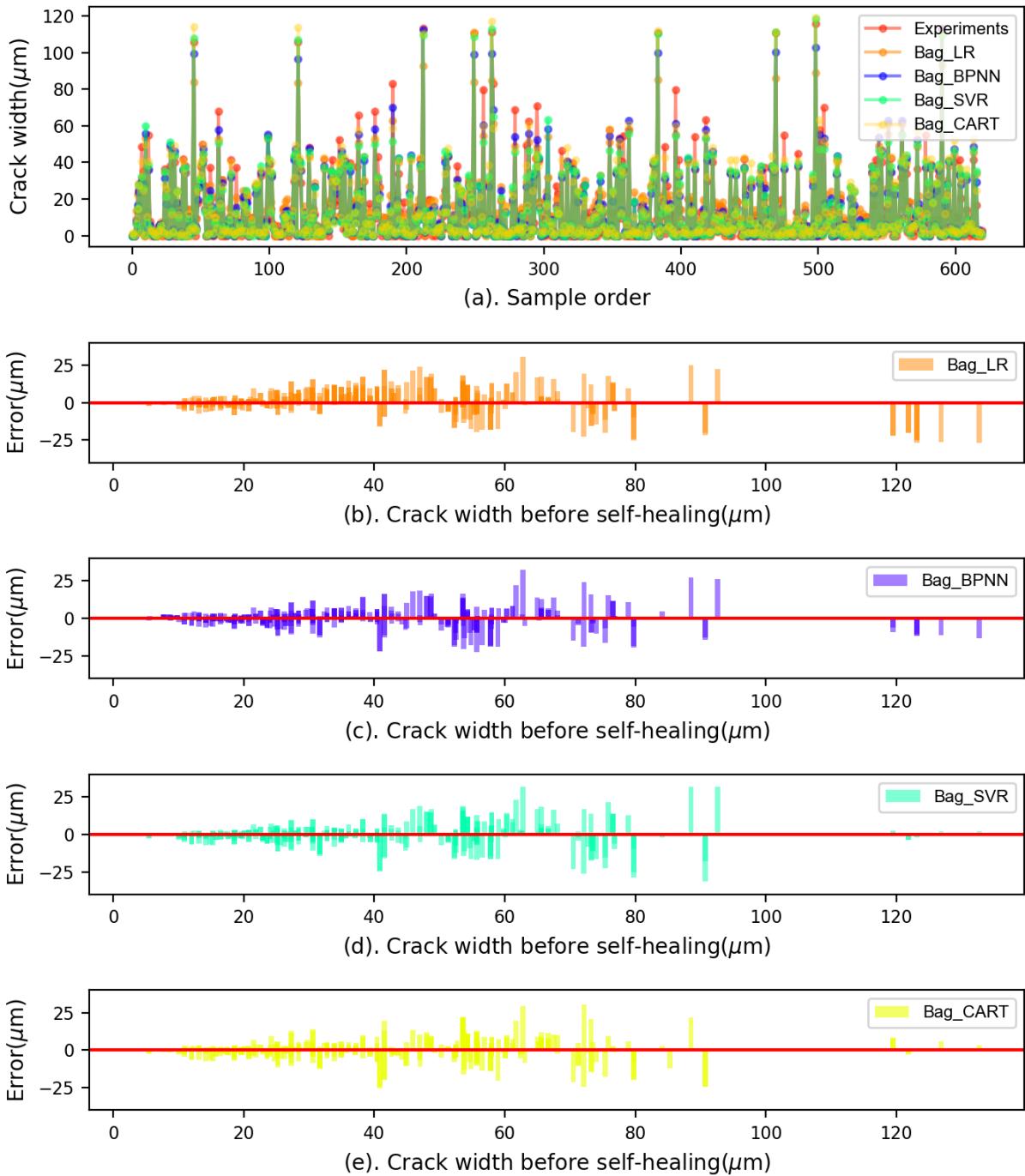


Figure 11: Comparison of predicted and observed crack width after self-healing of ECC for bagging ensemble models

not significant. Similar variations can be found when comparing SVR with Ada_SVR and bag_SVR.

After stacking, the error variations are much reduced and in the range between -20 and 20 μm as shown

in Figure 13 compared the range between -25 μm to 25 μm shown in Figure 10 to 12. More specifically, the

average absolute errors of stack_LR were 1.361, 4.932, 9.789 and 3.177 for initial cracks ranges less than 20,

between 20 to 60, between 60 to 100 and over 100 μm before self-healing, respectively. And the MAE value

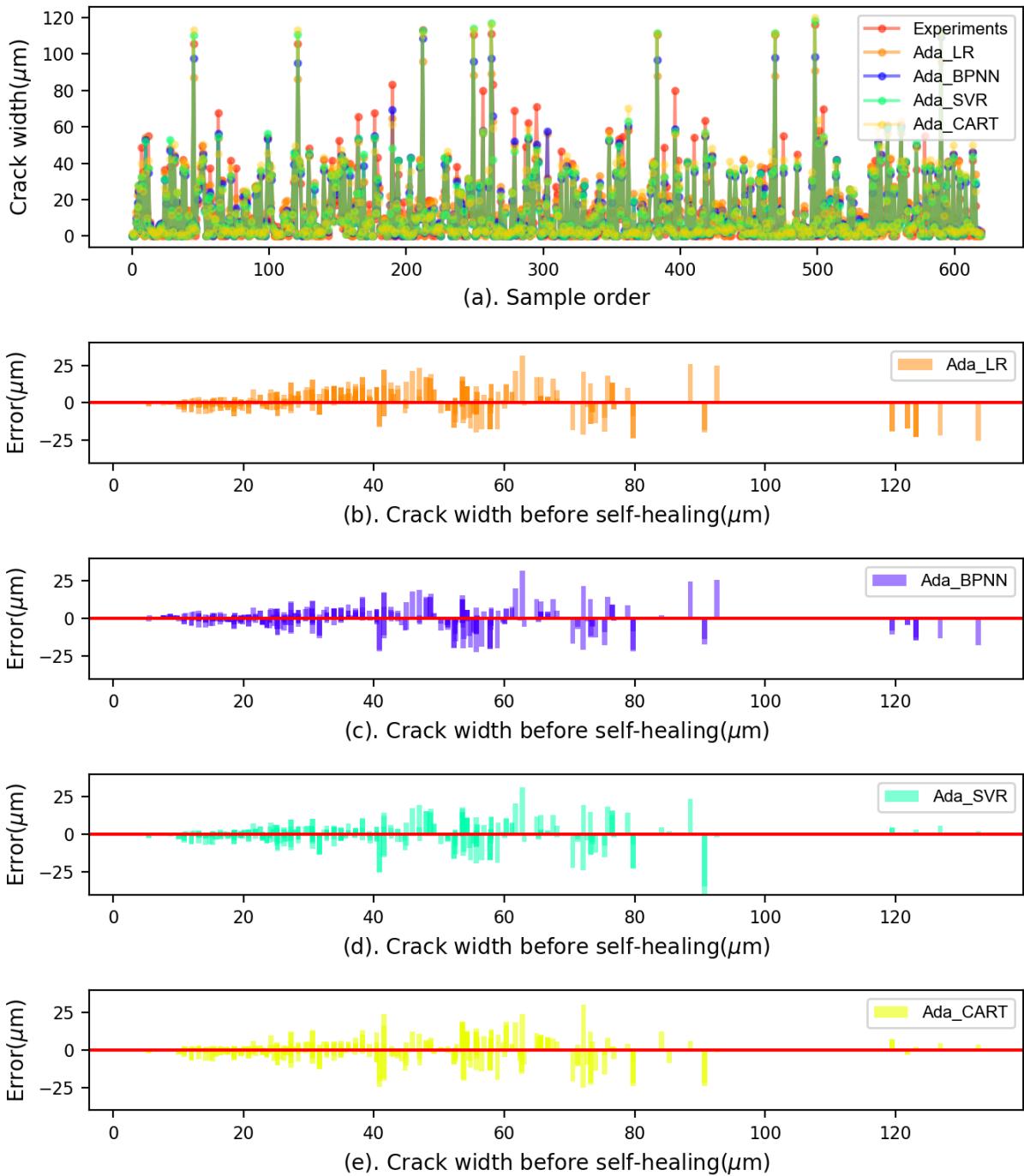


Figure 12: Comparison of predicted and observed crack width after self-healing of ECC for adaBoost ensemble models

for Stack_LR on whole dataset is 3.934 shown in Table 5. Therefore, the stack_LR model that combines multiple prediction models (BPNN, CART, SVR and LR) is superior to all other models shown the lowest variations.

It is known that smaller crack width is favourable for autogenous healing in concrete [64, 65] as small cracks consume less repair products to complete self-healing [66]. However, larger crack width will not heal

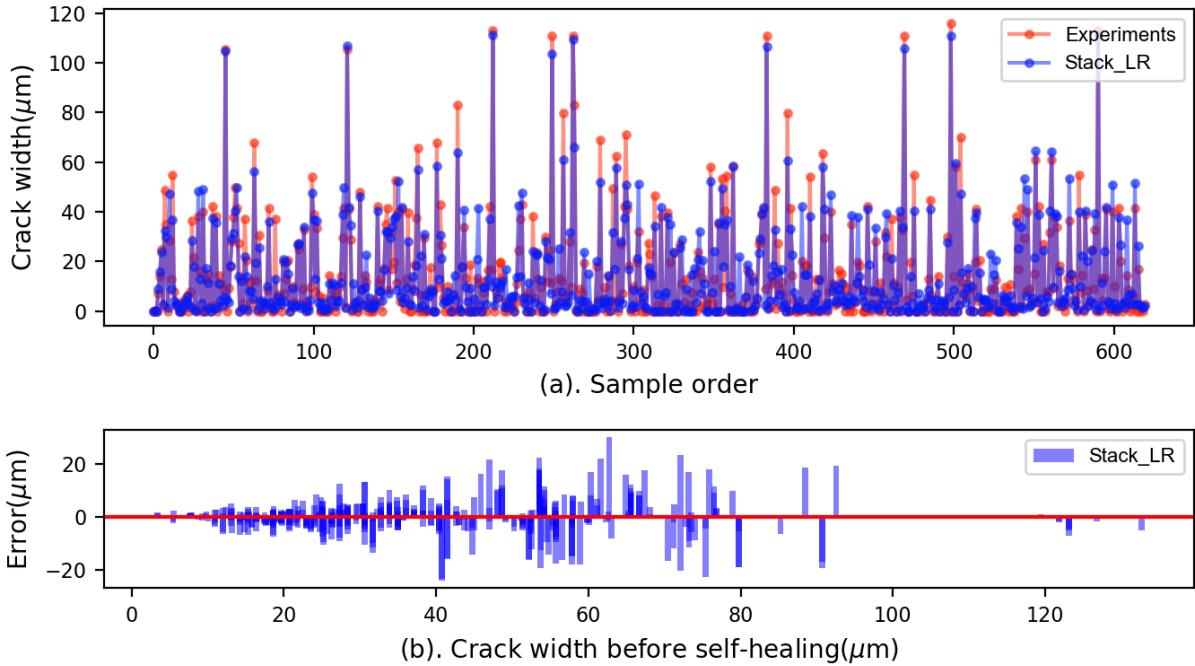


Figure 13: Comparison of predicted and observed crack width after self-healing of ECC for stacking ensemble models

452 completely or just heal partially. As shown in Figures b in 10 to 13, the variations between the observed and
 453 predicted results for the LR, bag_LR and Ada_LR models increased with the increase of crack width. For the
 454 initial crack width below 20 μm , the average absolute errors were less than 1.5, which is much lower than
 455 that for crack widths between 20 and 60 μm around 6.23, and between 60 and 100 μm around 10. And all
 456 other models had the same variation trend. Higher variations in predicted self-healing results compared to
 457 the observed results were discovered for larger initial crack widths (less than 100 μm). But the variations for
 458 those models were less than the LR, bag_LR and Ada_LR models. Specially, the LR, bag_LR and Ada_LR
 459 models showed much higher variations with average absolute error over 20 compared with other ML models
 460 with average absolute error less than 10 for the crack width over 100 μm .

461 Thus, it can be concluded that ensemble methods can significantly improve the prediction accuracy of
 462 individual methods. Besides, the stack_LR model that combines multiple prediction models (BPNN, CART,
 463 SVR and LR) is superior to all other models shown the lowest variations.

464 A box plot as shown in Figure 14 is created to show the distribution of RMSE results of each ML model
 465 from ten-fold cross validation. The RMSE values were calculated based on the differences between the
 466 predicted and observed crack widths. Box plot is a statistical tool that is used to depicting numerical data
 467 through their quartiles including maximum, minimum, median values of a dataset [67, 68].The medium
 468 value is shown as the red line within the box. The interquartile range (IQR) in each box covers the 50% (the
 469 lower 25% to the upper 75% quartiles) of the RMSE data point, while the whiskers drawn up and down to
 470 the maximum and minimum values represent 1.5 times the IQR from the RMSE data. All other points out of
 471 the whiskers range are outliers and shown as red dots. A mean value of RMSE equals to zero would indicate
 472 that the predictions perfectly fit the observed data. However, this is almost never achieved in practice [69].
 473 In general, the lower RMSE value the better the prediction performance of a model.

474 Assessment of the box plot revealed that the stack_LR model outperforms all other models because of
 475 its shortest IQR length and smallest RMSE values as shown in Figure 5. In contrast, the LR and bag_LR
 476 models have the longest IQR length and largest RMSE values, thereby suggesting that the LR model and its

477 ensemble methods are of low accuracy. Among the individual models, BPNN has the lowest RMSE, while
 478 SVR has the shortest IQR length but with three outliers (out of ten data points). In general, BPNN gave the
 479 most stable performance showing reasonable low RMSE values with short IQR length.

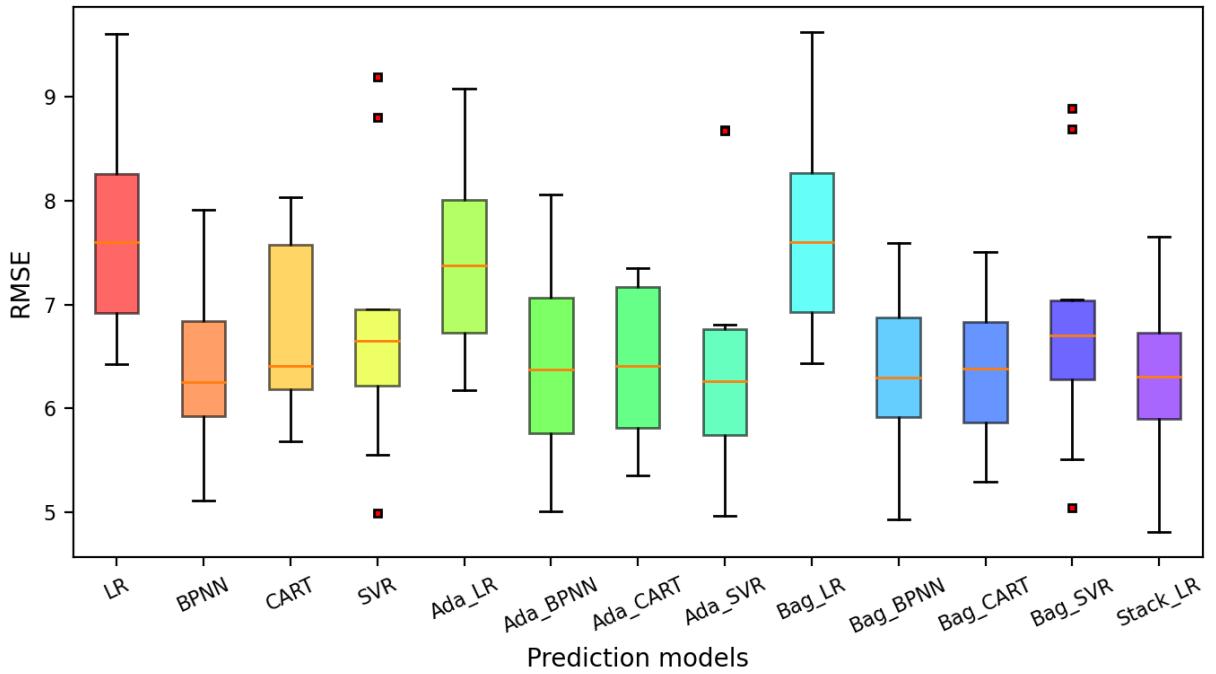


Figure 14: Ten-fold cross validation of RMSE by proposed ML models in prediction of self-healing ability of ECC

480 6 Conclusions

481 In this study, several individual and ensemble ML models were proposed to predict the self-healing ability of
 482 ECC. All the models were trained and validated based on the experimental results from nine ECC mixtures.
 483 Based on the results, the following conclusions can be drawn.

- 484 1. Among of the individual ML model studies, the BPNN model performed the best in terms of RMSE
 485 and R^2 .
- 486 2. All ensemble methods can generally improve the prediction accuracy of individual methods, however
 487 the improvement varies. It is found that Bagging method mainly enhanced the performance of BPNN
 488 and CART whereas AdaBoost method brought a considerable improvement for LR and SVR models.
- 489 3. Among all the ML models studied, the Stack_LR model demonstrated great prediction on self-healing
 490 of ECC and performed the best on MAE, RMSE and R^2 results. The assessment of the box plot also
 491 revealed that the stackLR model outperforms all other models because of its shortest IQR length and
 492 smallest RMSE values.
- 493 4. For the initial crack widths less than $60 \mu\text{m}$, the variations shown in SVR model are smaller than those
 494 presented in other models. However, the CART model showed smaller variations for the crack widths
 495 between 60 and $100 \mu\text{m}$ compared to the SVR and BPNN models. For crack widths larger than 100
 496 μm , the SVR model performed the best showing the smallest variations.

- 497 5. The computational results indicate that the individual and ensemble methods can be used to predict
 498 the self-healing ability of ECC. However, how to choose an appropriate base learner and ensemble
 499 method is critical. To improve the performance accuracy, researchers should employ different ensem-
 500 ble methods to compare their effectiveness with different ML models. The proposed individual and
 501 ensemble ML models can be used to predict the self-healing properties of ECC.
- 502 6. Future investigation and experimentation should be carried to extend the training dataset to include
 503 various crack width distributions and diverse influencing factors such as components, W/MC rate etc..
 504 In addition, more research should be undertaken to optimise parameters in ML models and develop a
 505 hybrid model to improve the prediction accuracy.

506 **References**

- 507 [1] Diane Gardner, Robert Lark, Tony Jefferson, and Robert Davies. A survey on problems encountered
 508 in current concrete construction and the potential benefits of self-healing cementitious materials. *Case*
 509 *studies in construction materials*, 8:238–247, 2018.
- 510 [2] E Cailleux and V Pollet. Investigations on the development of self-healing properties in protective
 511 coatings for concrete and repair mortars. In *Proceedings of the 2nd International Conference on Self-*
 512 *Healing Materials, Chicago, IL, USA*, volume 28, 2009.
- 513 [3] Ahmed Ramadan Suleiman and Moncef L Nehdi. Modeling self-healing of concrete using hybrid
 514 genetic algorithm–artificial neural network. *Materials*, 10(2):135, 2017.
- 515 [4] Waiching Tang, Omid Kardani, and Hongzhi Cui. Robust evaluation of self-healing efficiency in ce-
 516 mentitious materials—a review. *Construction and Building Materials*, 81:233–247, 2015.
- 517 [5] Carola Edvardsen. Water permeability and autogenous healing of cracks in concrete. *Materials Journal*,
 518 96(4):448–454, 1999.
- 519 [6] Waiching Tang, Guangwei Chen, and Shanyong Wang. Self-healing capability of ecc incorporating
 520 with different mineral additives - a review. In Changwen Miao, editor, *Proceedings of the 3rd Interna-*
 521 *tional RILEM Conference on Microstructre Related Durability of Cementitious Composites*, page 668,
 522 2016.
- 523 [7] Stefan Jacobsen, Jacques Marchand, and Hugues Hornain. Sem observations of the microstructure of
 524 frost deteriorated and self-healed concretes. *Cement and Concrete Research*, 25(8):1781–1790, 1995.
- 525 [8] Hans-Wolf Reinhardt and Martin Jooss. Permeability and self-healing of cracked concrete as a function
 526 of temperature and crack width. *Cement and concrete research*, 33(7):981–985, 2003.
- 527 [9] Mustafa Şahmaran and İ Özgür Yaman. Influence of transverse crack width on reinforcement corrosion
 528 initiation and propagation in mortar beams. *Canadian Journal of Civil Engineering*, 35(3):236–245,
 529 2008.
- 530 [10] CA Clear. The effects of autogenous healing upon the leakage of water through cracks in concrete.
 531 Technical report, 1985.
- 532 [11] Mustafa Sahmaran, Gurkan Yildirim, and Tahir K Erdem. Self-healing capability of cementitious
 533 composites incorporating different supplementary cementitious materials. *Cement and Concrete Com-*
 534 *posites*, 35(1):89–101, 2013.

- 535 [12] Toshiro Kamada and Victor C Li. The effects of surface preparation on the fracture behavior of
536 ecc/concrete repair system. *Cement and Concrete Composites*, 22(6):423–431, 2000.
- 537 [13] Victor C Li and Tetsushi Kanda. Innovations forum: engineered cementitious composites for structural
538 applications. *Journal of Materials in Civil Engineering*, 10(2):66–69, 1998.
- 539 [14] Erdoan zbay, Mustafa ahmaran, Mohamed Lachemi, and Hasan Erhan Yucel. Self-healing of
540 microcracks in high-volume fly-ash-incorporated engineered cementitious composites. *ACI Materials
541 Journal*, 110(1), 2013.
- 542 [15] Min Wu, Bj rn Johannesson, and Mette Geiker. A review: Self-healing in cementitious materials and
543 engineered cementitious composite as a self-healing material. *Construction and Building Materials*,
544 28(1):571–583, 2012.
- 545 [16] Haoliang Huang, Guang Ye, and Denis Damidot. Characterization and quantification of self-healing
546 behaviors of microcracks due to further hydration in cement paste. *Cement and Concrete Research*, 52:
547 71–81, 2013.
- 548 [17] Ahmed R Suleiman, Andrew J Nelson, and Moncef L Nehdi. Visualization and quantification of crack
549 self-healing in cement-based materials incorporating different minerals. *Cement and Concrete Com-
550 posites*, 103:49–58, 2019.
- 551 [18] J Zhou, S Qian, MG Sierra Beltran, G Ye, E Schlangen, and K van Breugel. Developing engineered
552 cementitious composite with local materials. In *International conference on microstructure related
553 durability of cementitious composites, Nanjing, China*, 2008.
- 554 [19] Victor C Li and En-Hua Yang. Self healing in concrete materials. In *Self healing materials*, pages
555 161–193. Springer, 2007.
- 556 [20] Zhigang Zhang, Shunzhi Qian, and Hui Ma. Investigating mechanical properties and self-healing be-
557 havior of micro-cracked ecc with different volume of fly ash. *Construction and Building Materials*, 52:
558 17–23, 2014.
- 559 [21] G rkan Yildirim, Mustafa Sahmaran, and Hemn Unis Ahmed. Influence of hydrated lime addition on
560 the self-healing capability of high-volume fly ash incorporated cementitious composites. *Journal of
561 Materials in Civil Engineering*, 27(6):04014187, 2014.
- 562 [22] Y Yang, M Lepech, and Victor C Li. Self-healing of ecc under cyclic wetting and drying. 2005.
- 563 [23] Mustafa Sahmaran, Mo Li, and Victor C Li. Transport properties of engineered cementitious compos-
564 ites under chloride exposure. *ACI Materials Journal*, 104(6):604, 2007.
- 565 [24] SZ Qian, J Zhou, and E Schlangen. Influence of curing condition and precracking time on the self-
566 healing behavior of engineered cementitious composites. *Cement and concrete composites*, 32(9):
567 686–693, 2010.
- 568 [25] Marai M Alshihri, Ahmed M Azmy, and Mousa S El-Bisy. Neural networks for predicting compressive
569 strength of structural light weight concrete. *Construction and Building Materials*, 23(6):2214–2219,
570 2009.
- 571 [26] Uwe Reuter, Ahmad Sultan, and Dirk S Reischl. A comparative study of machine learning approaches
572 for modeling concrete failure surfaces. *Advances in Engineering Software*, 116:67–79, 2018.

- 573 [27] Siamak Safarzadegan Gilan, Hamed Bahrami Jovein, and Ali Akbar Ramezanianpour. Hybrid sup-
574 port vector regression–particle swarm optimization for prediction of compressive strength and reft of
575 concretes containing metakaolin. *Construction and Building Materials*, 34:321–329, 2012.
- 576 [28] Fei Yan, Zhibin Lin, Xingyu Wang, Fardad Azarmi, and Konstantin Sobolev. Evaluation and prediction
577 of bond strength of gfrp-bar reinforced concrete using artificial neural network optimized with genetic
578 algorithm. *Composite Structures*, 161:441–452, 2017.
- 579 [29] Zaher Mundher Yaseen, Ravinesh C Deo, Ameer Hilal, Abbas M Abd, Laura Cornejo Bueno, Sancho
580 Salcedo-Sanz, and Moncef L Nehdi. Predicting compressive strength of lightweight foamed concrete
581 using extreme learning machine model. *Advances in Engineering Software*, 115:112–125, 2018.
- 582 [30] Kezhen Yan and Caijun Shi. Prediction of elastic modulus of normal and high strength concrete by
583 support vector machine. *Construction and Building Materials*, 24(8):1479–1485, 2010.
- 584 [31] Jui-Sheng Chou, Chih-Fong Tsai, Anh-Duc Pham, and Yu-Hsin Lu. Machine learning in concrete
585 strength simulations: Multi-nation data analytics. *Construction and Building Materials*, 73:771–780,
586 2014.
- 587 [32] Jafar Sobhani, Meysam Najimi, Ali Reza Pourkhorshidi, and Tayebeh Parhizkar. Prediction of the
588 compressive strength of no-slump concrete: A comparative study of regression, neural network and
589 anfis models. *Construction and Building Materials*, 24(5):709–718, 2010.
- 590 [33] Behzad Abounia Omran, Qian Chen, and Ruoyu Jin. Comparison of data mining techniques for pre-
591 dicting compressive strength of environmentally friendly concrete. *Journal of Computing in Civil
592 Engineering*, 30(6):04016029, 2016.
- 593 [34] Luthfi Muhammad Mauludin and Chahmi Oucif. Modeling of self-healing concrete: A review. *Jour-
594 nal of Applied and Computational Mechanics*, 5(Special Issue: Computational Methods for Material
595 Failure):526–539, 2019.
- 596 [35] Wenjun Wang, Nicolette G Moreau, Yingfang Yuan, Paul R Race, and Wei Pang. Towards machine
597 learning approaches for predicting the self-healing efficiency of materials. *Computational Materials
598 Science*, 168:180–187, 2019.
- 599 [36] M Chaitanya, P Manikandan, V Prem Kumar, S Elavenil, and V Vasugi. Prediction of self-healing char-
600 acteristics of ggbs admixed concrete using artificial neural network. In *Journal of Physics: Conference
601 Series*, volume 1716, page 012019. IOP Publishing, 2020.
- 602 [37] Xiaoying Zhuang and Shuai Zhou. The prediction of self-healing capacity of bacteria-based concrete
603 using machine learning approaches. *Computers, Materials & Continua* 59 (2019), Nr. 1, 2019.
- 604 [38] AS 3972-2010 General purpose and blended cements. URL [https://infostore.saiglobal.
605 com/en-au/standards/as-3972-2010-122323_saig_as_as_268436/](https://infostore.saiglobal.com/en-au/standards/as-3972-2010-122323_saig_as_as_268436/).
- 606 [39] Guangwei Chen. *Repeatability of self-healing in ECC with various mineral admixtures*. PhD thesis,
607 School of Architecture and Built Environment, University of Newcastle, Australia, 2021.
- 608 [40] John Neter, Michael H Kutner, Christopher J Nachtsheim, and William Wasserman. *Applied linear
609 statistical models*, volume 4. Irwin Chicago, 1996.
- 610 [41] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297,
611 1995.

- 612 [42] Vladimir Naumovich Vapnik. An overview of statistical learning theory. *IEEE transactions on neural*
 613 *networks*, 10(5):988–999, 1999.
- 614 [43] Xu Juncai, Ren Qingwen, and Shen Zhenzhong. Prediction of the strength of concrete radiation shield-
 615 ing based on ls-svm. *Annals of nuclear energy*, 85:296–300, 2015.
- 616 [44] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural*
 617 *processing letters*, 9(3):293–300, 1999.
- 618 [45] Yankun Li, Xueguang Shao, and Wensheng Cai. A consensus least squares support vector regression
 619 (ls-svr) for analysis of near-infrared spectra of plant samples. *Talanta*, 72(1):217–222, 2007.
- 620 [46] P Yuvaraj, A Ramachandra Murthy, Nagesh R Iyer, SK Sekar, and Pijush Samui. Support vector
 621 regression based models to predict fracture characteristics of high strength and ultra high strength
 622 concrete beams. *Engineering fracture mechanics*, 98:29–43, 2013.
- 623 [47] Alex J Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and comput-*
 624 *ing*, 14(3):199–222, 2004.
- 625 [48] Abhijit Mukherjee and Sudip Nag Biswas. Artificial neural networks in prediction of mechanical
 626 behavior of concrete at high temperature. *Nuclear engineering and design*, 178(1):1–11, 1997.
- 627 [49] Hosein Naderpour, Amir Hossein Rafiean, and Pouyan Fakharian. Compressive strength prediction of
 628 environmentally friendly concrete using artificial neural networks. *Journal of Building Engineering*,
 629 16:213–219, 2018.
- 630 [50] Leo Breiman. *Classification and regression trees*. 2017.
- 631 [51] S Dan and P Colla. Cart: Tree-structured non-parametric data analysis. *Salford Systems, San Diego*,
 632 CA, 1995.
- 633 [52] R Put, C Perrin, F Questier, D Coomans, DL Massart, and Y Vander Heyden. Classification and
 634 regression tree analysis for molecular descriptor selection and retention prediction in chromatographic
 635 quantitative structure–retention relationship studies. *Journal of Chromatography A*, 988(2):261–276,
 636 2003.
- 637 [53] Dimitrios Frosyniotis, Andreas Stafylopatis, and Aristidis Likas. A divide-and-conquer method for
 638 multi-net classifiers. *Pattern Analysis & Applications*, 6(1):32–40, 2003.
- 639 [54] Thomas G Dietterich. Ensemble methods in machine learning. In *International workshop on multiple*
 640 *classifier systems*, pages 1–15. Springer, 2000.
- 641 [55] Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.
- 642 [56] Yoav Freund, Robert E Schapire, et al. Experiments with a new boosting algorithm. In *Icml*, volume 96,
 643 pages 148–156. Citeseer, 1996.
- 644 [57] Sebastian Raschka. Mlxtend: Providing machine learning and data science utilities and extensions to
 645 python’s scientific computing stack. *The Journal of Open Source Software*, 3(24), 2018. doi: 10.21105/
 646 joss.00638. URL <http://joss.theoj.org/papers/10.21105/joss.00638>.
- 647 [58] Joseph Sill, Gábor Takács, Lester Mackey, and David Lin. Feature-weighted linear stacking. *arXiv*
 648 *preprint arXiv:0911.0460*, 2009.

- 649 [59] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selec-
650 tion. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995.
- 651 [60] Ofer Dor and Yaoqi Zhou. Achieving 80% ten-fold cross-validated accuracy for secondary structure
652 prediction by large-scale training. *Proteins: Structure, Function, and Bioinformatics*, 66(4):838–845,
653 2007.
- 654 [61] Thierry Vidal, Arnaud Castel, and Raoul François. Analyzing crack width to predict corrosion in
655 reinforced concrete. *Cement and concrete research*, 34(1):165–174, 2004.
- 656 [62] Kejin Wang, Daniel C Jansen, Surendra P Shah, and Alan F Karr. Permeability study of cracked
657 concrete. *Cement and concrete research*, 27(3):381–393, 1997.
- 658 [63] Jishen Qiu, Han Siang Tan, and En-Hua Yang. Coupled effects of crack width, slag content, and
659 conditioning alkalinity on autogenous healing of engineered cementitious composites. *Cement and*
660 *Concrete Composites*, 73:203–212, 2016.
- 661 [64] Emily N. Herbert and Victor C. Li. Self-Healing of Microcracks in Engineered Cementitious Compos-
662 ites (ECC) Under a Natural Environment. *Materials*, 6:2831–2845, 2013. doi: 10.3390/ma6072831.
- 663 [65] Hezhi Liu, Qian Zhang, Chongshi Gu, Huaizhi Su, and Victor Li. Influence of microcrack self-healing
664 behavior on the permeability of engineered cementitious composites. *Cement and Concrete Compos-*
665 *ites*, 82:14–22, 2017.
- 666 [66] Nele De Belie, Elke Gruyaert, Abir Al-Tabbaa, Paola Antonaci, Cornelia Baera, Diana Bajare, Aveline
667 Darquennes, Robert Davies, Liberato Ferrara, Tony Jefferson, et al. A review of self-healing concrete
668 for damage management of structures. *Advanced Materials Interfaces*, 5(17):1800074, 2018.
- 669 [67] Woubishet Zewdu Taffese, Esko Sistonen, and Jari Puttonen. Caprm: Carbonation prediction model
670 for reinforced concrete using machine learning methods. *Construction and Building Materials*, 100:
671 70–82, 2015.
- 672 [68] Oladimeji B Olalusi and Panagiotis Spyridis. Machine learning-based models for the concrete breakout
673 capacity prediction of single anchors in shear. *Advances in Engineering Software*, 147:102832, 2020.
- 674 [69] Wikipedia contributors. Root-mean-square deviation — Wikipedia, the free encyclopedia, 2021.
675 URL https://en.wikipedia.org/wiki/Root-mean-square_deviation. [Online;
676 accessed 28-August-2021].