AI-driven formulator skill augmentation
- Co-formulation of Human and AI for cosmetic development -

Arai, Toshihiro[1*]; Ohara, Kaho[1]; Kimura, Yuki[1]; Hase, Takeshi[2]; Sathiyananthavel, Mayuri[2]; Matsuoka, Yukiko[2]; Yachie, Ayako[2]

[1] Frontier Research Center, POLA Chemical Industries, Inc., Kanagawa, Japan; [2] SBX Corporation, Tokyo, Japan

* Arai Toshihiro, 560 Kashio-cho Totsuka-ku Yokohama Japan, +81-45-826-7231

Abstract

In the cosmetics industry, quickly delivering new products for the constantly diversifying and evolving market is essential. To quickly develop products that meet these needs, formulators are required to improve their ability to design novel formulations effectively. On the other hand, the current formulation design practice in repeating hypothesis construction and prototype-based verification experiments is time-consuming and resource intensive. Therefore, we considered how to build an AI that can be trained using cosmetics information to accelerate the development of the formulator. To accomplish this, we designed an AI that predicts the quality of the finished cosmetics and provides appropriate information for formulators. First, we acquired data related to cosmetic products such as ingredients, creation process, physical properties or tactile sensations. Then we developed prediction models that connects the tactile sensation data with formulation data using a machine learning algorithm, and designed an AI system that uses these prediction models. Through use of the AI, formulators were able to experience the improvement of their design skill efficiency. By expanding the limits of human skills with AI, cosmetic chemists should be able to develop innovative and revolutionary formulations and give cosmetics possibilities of new functions and uses.

Keywords

Skill Augmentation; Artificial Intelligent; Tactile Sensation; Formulation; Product Development

Introduction

In the cosmetics industry, customer's needs are diversifying and evolving so formulators need to efficiently design various formulations more than ever before. Additionally, sustainability is also an important theme in the cosmetics industry and companies are starting research and implementation of sustainability. Specifically, it is crucial for formulators to create cosmetics using only environmentally friendly ingredients, based on each country's increasing restrictions of chemical substances such as the European Green Deal strategy [1]. However, the outcome of this initiative may result in development being slowed as the variety of ingredients is insufficient to meet market needs, and it is additionally likely to hinder improvement of product quality. In other words, formulators need to quickly and effectively combine environmentally friendly ingredients to efficiently design formulations that meet various customer needs.

In order to quickly design a wide variety of formulations, formulators should improve not only their knowledge of ingredients and emulsions but also their skill to design formulations by considering the impact on the quality of cosmetics when specific ingredients are combined. For improving this "design skill", formulators usually train themselves by repeating hypothesis construction and verification by creating prototypes. However, this trial and error process of prototype creation takes a lot of time and it is not an efficient way to improve their "design skill". On the other hand, there are few reported systems for formulator training and only programs that teach basic knowledge about formulation are available [2].

Therefore, in order to effectively improve "design skill" of formulators, we considered how to build a training cycle with an AI that can be trained using cosmetic information. Since the AI is able to very rapidly predict the quality of the finished product from the formulation, formulators can nearly instantly compose and verify their hypotheses *in silico*. In addition, information regarding the relationship between ingredients and cosmetic qualities discovered by the AI will broaden formulator's knowledge. Furthermore, by creating actual prototypes after AI simulation and evaluating them with comparing collected data, we expect useful feedback can be obtained and the next formulation design can be precisely implemented.

To achieve this, we developed an AI that predicts the quality of finished cosmetic products and provides information such as the ingredient and the creation process related to cosmetic qualities to formulators. In this study, we focused on physical property and tactile sensation as cosmetic qualities, which are crucial for formulation design. First, we collected data on the in-house marketed cosmetics that can be used for machine learning model (ML model) training. Input data of ML models is mainly formulation information such as ingredients and creation process that have been designed by formulators. Additionally, physical properties of formulation are thought to be important for predicting the tactile sensation of a cosmetic product because there are several studies on the relationship between tactile sensation and physical property of cosmetics [3, 4, 5, 6]. However, physical properties are not suitable as input information because they are known only after the cosmetic product is made. Therefore, we decided to develop a multi-step prediction process. Using the data of our in-house marketed cosmetics, we first created ML models that predict the physical properties of the finished product based on information of the ingredients and formulation procedures. Then we created ML models that predict tactile sensations of the finished product based on physical property information. Finally, we developed an AI system equipped with those ML models and some formulators evaluated whether the AI was effective for improvement of their "design skill".

Materials and Methods

Overview

The following is an overview of the AI development for cosmetics (Figure 1). Step1: Data containing features and objective values was generated by collecting and/or measuring information on formulation, physical properties, and tactile sensation scores of the target formulation. In data pre-processing, we removed several features that were unnecessary to predict objective values (e.g., physical properties and tactile sensations). Step 2: We performed feature selection analyses on the pre-processed data to select useful features to predict objective values. For data augmentation of training data, we used a state-of-the-art generative model to generate synthetic datasets that preserve characteristics of original real datasets. By using the resultant training datasets, we trained machine learning (ML) models to predict objective values with cross validation. Then, we selected the best ML models for each physical property and tactile sensation. Step 3: Finally, we designed and implemented an AI system that can execute the ML models on a local PC and be used by formulators. Specific details of the implementation are described below.

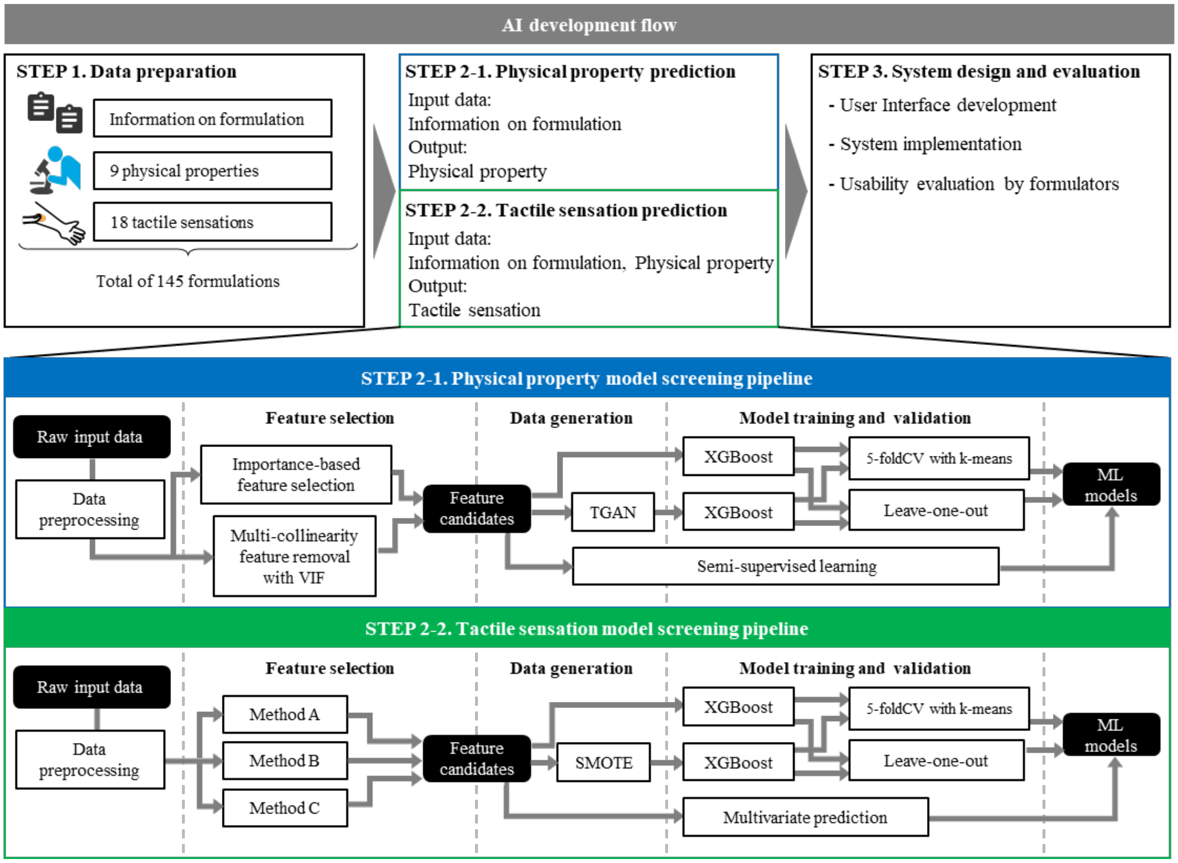

Figure 1 Flowchart of AI development

Step 1. Data preparation

We collected 145 in-house marketed skincare products such as toners, milky lotions, creams, and gels. Information on formulation was collected from our database. Details of the information on formulation are described below.

- Ingredients: Information on the ingredients contained in each formulation.
- Formulation Properties: Information on the phase state of the formulation (oil-in-water, water-in-oil, etc.).
- Creation process of formulation: Information such as temperature and stirring conditions during emulsification.
- Physical Properties: Physical properties that are commonly used for cosmetic evaluation were obtained. We obtained data for 9 types of physical properties such as pH, specific gravity, viscosity at shear rate of 100 s^{-1}, thermal conductivity, coefficient of friction when applied on artificial leather, peel force, residue on drying, and contact angle. In this step

some characteristics of the formulation made it impossible to measure specific physical properties (e.g., pH cannot be measured in the water-in-oil formulation). Hence, such cases were excluded from the acquisition and therefore not included in the training data.

- Tactile Sensations: To build a ML model, a large amount of tactile sensation data that can be compared between multiple cosmetic formulations was needed. Data on the strength of 18 commonly used tactile sensation descriptors was obtained using the Check All that Apply (CATA) Method, which does not require training and can be done in a short time [7, 8]. The 18 tactile sensation descriptors are as follows: slippery; sticky; thick; moist; soft; hard; warm; cold; rough; smooth; wearability; spreadable; becomes firm; light; watery; absorbable; coating; and oily. In the tactile sensation data, since few formulations were recognized as "rough" or "warm" and most of the formulations had those at an intensity of 0, we used the numerical data for these items as a binary value 0 or 1.

Step 2. ML model screening

Preprocessing

For numerical features (e.g., the ingredient information), we used the blending ratio of each ingredient (%). For categorical data including creation process information (e.g., blending type), we used one-hot vector encoding to represent the features of the categorical data. For example, for a formulation with a blending type, the value of the binary variable of the formulation is equal to 1 while that for a formulation without the blending type is equal to 0. Then, in processing of the data, ingredients that were not considered to affect the physical properties such as fragrances and extracts in terms of blending ratio or application, were removed from the formulation information. In addition, since ingredient distributions were different between aqueous formulation (toners or gels) and emulsion formulation (milky lotions or creams), we built two different ML models for the two different types of formulations.

Feature selection

We selected useful features for prediction of physical properties, based on statistical and network analyses of training data. First, the correlation coefficients between feature candidates and physical properties were calculated. Then, we performed network analysis with the Maximum Spanning Tree Algorithm [9] to investigate the relationship between features and physical properties. Finally, we calculated feature weights for each feature by using the Relief algorithm [10]. The total feature importance was determined by considering the outputs from these three analyses. We selected top 10 or 20 features with highest feature importance.

In addition, we calculated variable inflation factor (VIF), which indicates the influence of multicollinearity, of each feature to remove features with multicollinearity, i.e., we removed features with the value of VIF greater than a given threshold value.

For the tactile sensation prediction, the same procedure as the physical property model was used for feature selection. In addition to the feature selection, to examine the impact of formulation information on prediction accuracy, we considered the following three methods for building the tactile sensation models. Method A was to train a prediction model using only physical properties as feature values. Method B was to train a prediction model after feature selection for physical properties and formulation information. For Method C only the feature selection of formulation information was performed, and then a ML model was trained by using the selected features and all physical property values.

Data generation

The amount of data was not large enough to train models with high accuracy, because it was difficult to measure the physical properties of many cosmetics in a short period of time. In order to address this issue, we took three computational approaches. (i) We used a state-of-the-art data generation algorithm, tabular-data generative adversarial network (TGAN) [11], to generate synthetic data from a small amount of measured data, and used the synthetic data together with real training data to train ML models. (ii) In order to train ML models, we used semi-supervised learning (SSL) algorithms, i.e., SSL algorithms increase the amount of training data, by assigning labels (e.g., physical property values) for unlabeled data (e.g., formulation data of in-house cosmetics for which physical property values were not measured) by using pre-trained ML models. (iii) Synthetic Minority Over-sampling Technique (SMOTE) [12] was applied for oversampling training data with binary objective values (e.g., “rough” and “warm”).

Modeling

We used a state-of-the-art machine learning algorithm, XGBoost, to train prediction models. The XGBoost is based on a gradient tree boosting algorithm that trains and updates a large number of weak learners in boosting steps and builds a strong learner that integrates them as an ensemble learner [13].

We built the ML models to predict each physical property, based on the XGBoost algorithm and features selected by using feature selection methods (see “Feature selection” for details). We also built the ML models with the SSL algorithms or synthetic data generated by TGAN (see “Data generation” for details).

We also used the XGBoost algorithm to build ML models for tactile sensation prediction. Since it was assumed that the prediction accuracy could be improved by considering the interaction between tactile sensations, we used a “Multivariate Prediction [14]” approach, which uses the predicted tactile sensation score as feature to predict other tactile sensations.

Since the values of tactile sensations, “rough” and “warm”, were binary values (e.g., 0 or 1), we built binary classification models based on XGBoost to predict the tactile sensations.

Validation

We performed 5-fold cross validation (CV5), with 80% of the data being training data and 20% of the data being test data, to evaluate predictive performances of the ML models. In the evaluation processes, the small amount of data may induce bias in the distribution of the training data and the validation data, resulting in inappropriate performance evaluation. In order to address this issue, we used K-means to equally divide formulations in each cluster into the 5-fold datasets. On the other hand, leave-one-out cross validation (LOO) was also performed to evaluate the ML models to compare the results of CV5.

We used two evaluation metrics, Normalized Root Mean Square Error (NRMSE) and R^2 score to evaluate predictive performances of the ML models. Since the values of NRMSE are strongly influenced by outliers and may not correlate with intuitive prediction accuracy, we evaluated the ML models based on both of NRMSE and R^2 scores, though NRMSE is a standard statistical measure to evaluate predictive performance of regression models. We used Cohen’s Kappa coefficient that represents reproducibility as an evaluation metric for binary classification model. Furthermore, we evaluated the validity of the predictions for the training and validation data based on observed-predicted plots. In the observed-predicted plots, horizontal and vertical axis denotes the observed and predicted value, respectively. If there are a larger number of plots near the diagonal, the prediction is thought to be highly accurate and versatile.

Step 3. System design and evaluation

Combining the in-house database and created ML models, we designed an AI system. The system pipeline consisted of extracting features from the input formulation information, executing the calculation of the physical property ML models, and then further executing the calculation of the tactile sensation ML models to output the final prediction results. We also evaluated the effectiveness of the AI by having formulators (n=4) use the AI in their development work for about 3 months. Then we conducted questionnaire to ask the formulators the effect of the AI for "design skill" and efficiency of formulation design.

Results

Physical property ML model evaluation

We evaluated predictive performance of the ML models. Table 1 shows the values of NRMSE and those of R^2 for the best ML models to predict physical property. For each physical property, the best ML models show high predictive performance, i.e., the value of NRMSE and that of R^2 for the best ML models are less than 0.250 and greater than 0.600, respectively.

Table 1 Prediction accuracy of each physical property ML model

Type: Aqueous indicates the models trained with the data of aqueous formulation (toners and gels). Type: Emulsion indicates the models trained with the data of emulsion formulation (milky lotions and creams). Conditions represent the used algorithm, feature selection and validation method.

Type	Objective	NRMSE	R^2	Conditions
Aqueous	Contact angle	0.059	0.863	XGBoost, VIF, LOO
	Peel force	0.584	0.635	SSL, VIF
	pH	0.038	0.866	XGBoost, Top20, CV5
	Residue	0.099	0.773	XGBoost, Top20, LOO
	Sheer stress	0.609	0.876	SSL, VIF
	Specific gravity	0.007	0.819	XGBoost, Top10, CV5
	Thermal conductivity	0.018	0.845	XGBoost, Top20, LOO
	Viscosity	0.592	0.627	SSL, VIF
Emulsion	Contact angle	0.059	0.863	XGBoost, VIF, LOO
	Peel force	0.129	0.836	XGBoost, Top10, CV5
	pH	0.055	0.516	XGBoost, Top20, CV5
	Residue	0.098	0.960	XGBoost, VIF, CV5
	Sheer stress	0.044	0.935	SSL, VIF
	Specific gravity	0.014	0.623	XGBoost, VIF, CV5
	Thermal conductivity	0.230	0.950	SSL, Top20
	Coefficient of friction	0.075	0.511	XGBoost, VIF, CV5
	Viscosity	0.513	0.721	SSL, Top20

Furthermore, as shown in an observed-predicted plot for a representative physical property (Residue) in Figure 2, the measured and predicted values for each formulation are similar to each other and the plots are aligned diagonally. These results indicate that the best ML models are versatile models to accurately predit various physical properties of cosmetic formulation.

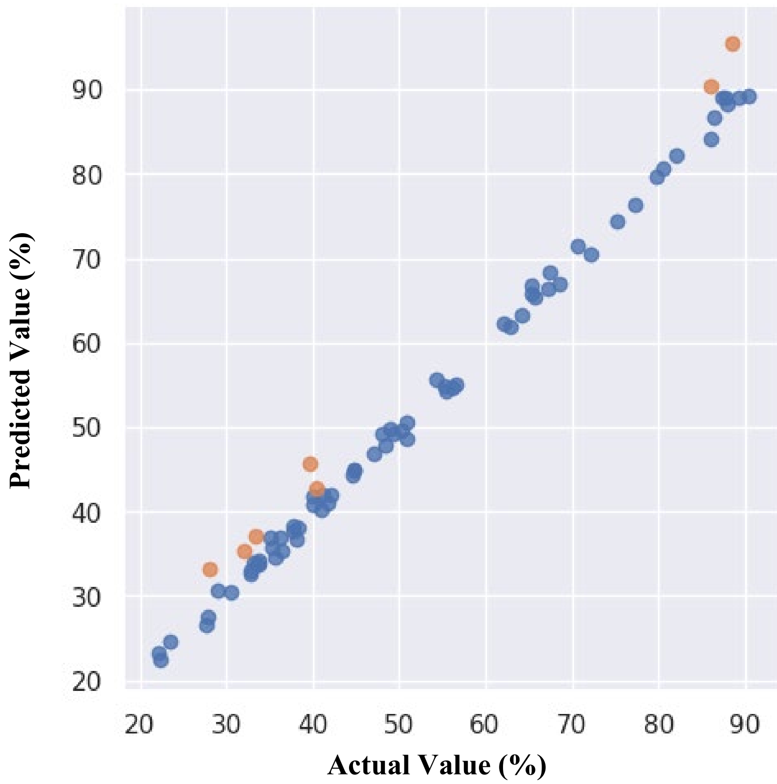

Figure 2 Observed-predicted plot for residue ML model of emulsion formulation

The blue and orage plots represent the results for the training data and those for the validation data, respectively.

Tactile sensation ML model evaluation

For each tactile sensation, the best ML models show high predictive performance (see Table 2), i.e., the values of Kappa score for warm and rough are 1.000, while the values of NRMSE and those of R^2 for the other tactile sensation are less than 0.250 and greater than 0.600, respectively. Furthermore, as shown in an observed-predicted plot for a representative tactile sensation, “moist”, in Figure 3, the measured and predicted values for each formulation are similar to each other and the plots are aligned diagonally. These results indicate that these best ML models listed in Table 2 are also versatile to accurately predict various tactile sensations (except “becomes firm” of emulsion and see “Discussion” about this result).

Table 2 Prediction accuracy of each tactile sensation ML models

Type: Aqueous indicates the models trained with the data of aqueous formulation (toners and gels). Type: Emulsion indicates the models trained with the data of emulsion formulation (milky lotions and creams). Conditions represent the used algorithm, feature selection and validation method.

Type	Objective	NRMSE	R^2	Kappa	Conditions
Aqueous	Hard	0.579	0.781		Multivariate, Method C
	Slippery	0.283	0.910		XGBoost, Method B, CV5
	Moist	0.142	0.775		XGBoost, Method C, CV5
	Thick	0.181	0.892		XGBoost, Method A, LOO
	Wearability	0.168	0.431		XGBoost with TGAN, Method C, LOO
	Smooth	0.143	0.595		Multivariate, Method B
	Spreadable	0.124	0.795		XGBoost, Method B, LOO
	Becomes firm	0.276	0.743		XGBoost, Method A, LOO
	Light	0.204	0.628		Multivariate, Method C
	Sticky	0.381	0.797		XGBoost, Method C, CV5
	Watery	0.237	0.848		XGBoost, Method C, CV5
	Soft	0.250	0.560		XGBoost, Method C, CV5
	Cold	0.257	0.640		XGBoost, Method C, LOO
	Absorbable	0.237	0.614		XGBoost, Method A, CV5
	Coating	0.245	0.670		Multivariate, Method C
	Oily	0.253	0.862		Multivariate, Method B
	Rough			1.000	XGBoost with SMOTE, Method B, CV3
	Warm			1.000	XGBoost with SMOTE, Method B, CV3
Emulsion	Hard	0.531	0.720		Multivariate, Method B
	Slippery	0.218	0.674		XGBoost, Method B, CV5
	Moist	0.094	0.498		XGBoost, Method A, LOO
	Thick	0.325	0.604		Multivariate, Method B
	Wearability	0.078	0.609		XGBoost, Method C, LOO
	Smooth	0.164	0.408		Multivariate, Method B
	Spreadable	0.150	0.401		Multivariate, Method B
	Becomes firm	0.424	0.389		XGBoost, Method B, CV5
	Light	0.247	0.414		XGBoost, Method C, CV5
	Sticky	0.352	0.743		Multivariate, Method B
	Watery	0.269	0.703		Multivariate, Method B
	Soft	0.133	0.646		XGBoost, Method A, CV5
	Cold	0.328	0.672		Multivariate, Method B
	Absorbable	0.164	0.601		Multivariate, Method B
	Coating	0.280	0.737		Multivariate, Method B
	Oily	0.318	0.766		XGBoost, Method B, CV5
	Rough			1.000	XGBoost, Method A, CV3
	Warm			1.000	XGBoost, Method C, CV3

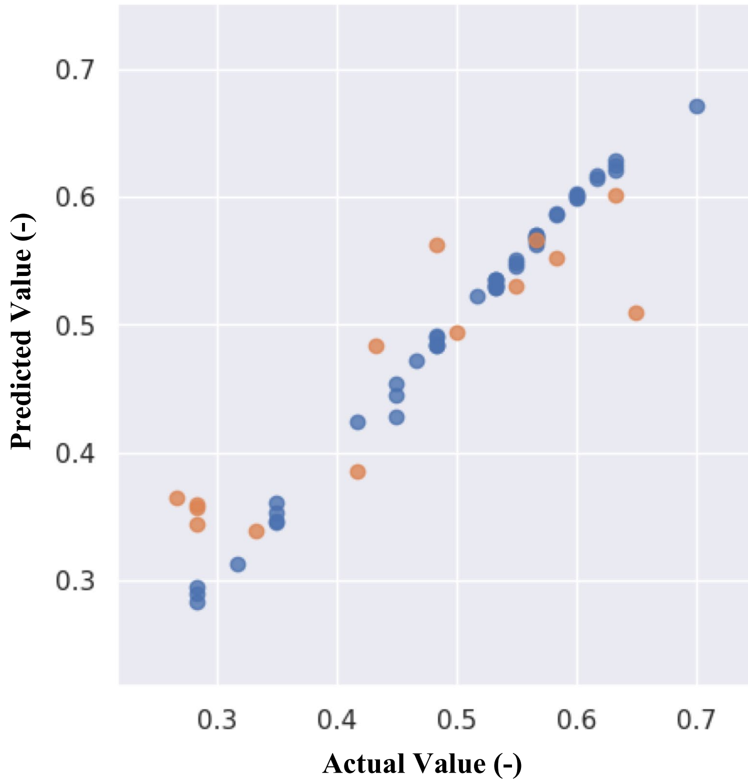

Figure 3 Observed-predicted plot for "moist" ML model of aqueous formulation

The blue and orage plots represent the results for the training data and those for the validation data, respectively.

AI system design and evaluation

By combining the created ML models and formulation database, we constructed a system that displays the prediction results: physical properties; tactile sensations; and the feature information that contributes to physical properties and tactile sensations (Figure 4). In addition, according to the data, it is possible to identify similar formulations as reference information, and to check the ingredients used in the formulation. The system requires only formulation and creation process information as input, and automatically executes the ML models internally.

Then the AI was used by several formulators (n=4). After about three months of AI trials, we conducted questionnaire and obtained some answers from formulators as shown in Table 3. All the formulators agreed or partly agreed on the effectiveness of the AI.

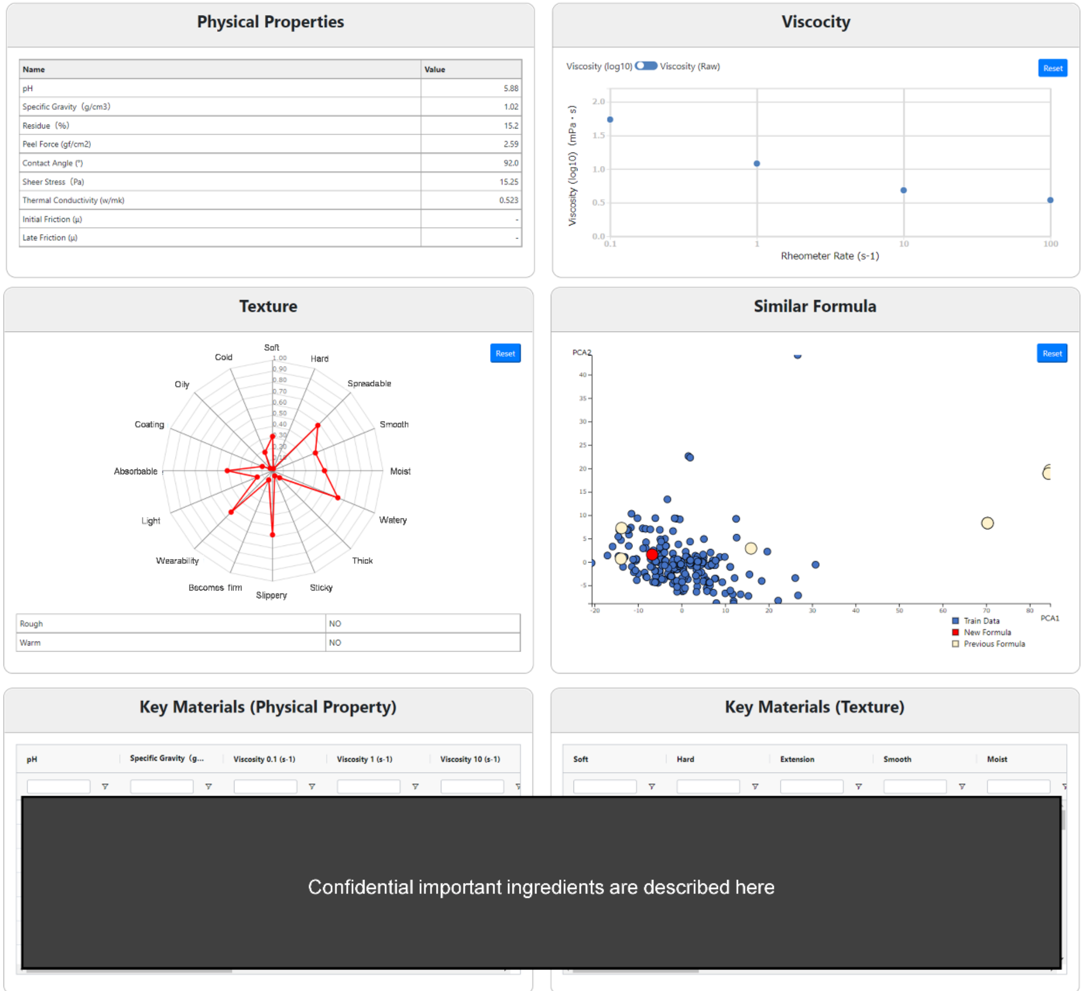

Figure 4 Output screen of developed AI

Table 3 Responses of questionnaire from formulators

Questionnaire	Response	Number	Comments
Did you gain any insights from using the AI and the data?	Agree	1	- I learned about new ingredients that affect tactile sensation and was able to use them effectively to create a new formulation.
	Partly agree	3	- The AI can be used to provide guidance to junior formulator that does not rely solely on my experience. - The tactile sensation data was useful for choosing base formulations. - For the beginning of formulation design, the AI was useful for learning the effect of ingredients on formulation qualities.
	Disagree	0	None
Did the AI make your formulation design more efficient?	Agree	2	- By entering candidate formulations, tactile sensation can be predicted and formulation strategies can be developed efficiently. - Preliminary simulation helped to decrease the number of formulations to be made.
	Partly agree	2	- Physical property prediction was useful for formulation design. - The AI and the data can be useful, but only for the initial formulation design.
	Disagree	0	None

Discussion

Marketed cosmetic products contain not only base ingredients but also multiple ingredients such as extracts and additives for quality control. When using information on these formulations for modeling, it is important to correctly identify the features that are effective for prediction of physical properties and tactile sensations. Otherwise, the accuracy of the created ML model would be likely low. In fact, through the model screening, it was confirmed that additives such as extracts that are included only in formulations that have characteristic physical properties were used as important features by the created model. Additionally, it led to decreasing the accuracy of the ML model. Therefore, it is effective for feature candidates that are included as appealing ingredients in a particular formulation to be considered as noise and removed in advance.

In addition, the ingredients used in aqueous formulation (toners, gels) and emulsion formulation (milky lotions, creams,) were very different, and the number of features contributing to the physical properties was also different. Therefore, if the information on these formulations was used as training data without separating it, the feature values that contribute to the physical properties could not be correctly learned, and prediction accuracy of the ML model tended to be low. It is considered effective to carefully examine the features that contribute to the physical properties for each cosmetic formulation and create a ML model as a separate data set based on these features.

For physical property prediction, the appropriate feature selection method differed depending on the physical properties to be predicted, and it was necessary to combine each method appropriately. In addition, when the data range of actual measured values was unevenly distributed, the prediction accuracy for a particular physical property range likely become low. Hence it is effective to add training data with a well distributed data range. In this study, we applied TGAN or SSL as data generation methods. In fact, prediction accuracy was better when using SSL as the data generation method than TGAN. This result suggests that this is because SSL more easily generates data that is closer to the actual data, and thus more useful information for training can be obtained.

As a result of ML model creation for tactile sensation prediction, the prediction accuracy was improved by adding the physical property values as features. Therefore, it was found that using physical property values as features is effective for the prediction of the tactile sensation. On the other hand, some tactile sensation descriptors with low prediction accuracy were found even when the physical property values were included as features. This is because the number of physical property values that can be obtained is limited, and not all physical characteristics that affect tactile sensation can be measured. On the other hand, some research has shown that there is a relationship between tactile sensation descriptors [15, 16]. Therefore, we used Multivariate Prediction, in which one predicted tactile sensation score is used as the feature to predict other tactile sensation score. This method is based on the assumption that the effect of the lack of features is greater than the effect of multicollinearity, although the use of predicted values for the prediction of other tactile sensation may increase the effect of multicollinearity. In fact, this method improved the prediction accuracy for some of the tactile sensations and was able to predict the tactile sensations with sufficient accuracy. This suggests that taking into account the interaction between tactile sensations in the prediction is a key point for accurate prediction.

On the other hand, prediction accuracy of the ML model of “becomes firm” for emulsion formulation was not so high. Although several methods were applied, the ML models of “becomes firm” tended to be overfitting and the prediction accuracy for test data became low. Since especially the prediction for specific formulations was not accurate, to analyze the characteristic of those formulations and add actual data on the similar formulations would be effective for improving the prediction accuracy. Further data collection is being conducted.

In this study, the ML models were created using data on in-house cosmetics for which the ingredients and creation processes are known. Although we did our best to remove the features that would become noise in the prediction, we cannot deny that there are cases in which predictions are made with reference to ingredient characteristics unique to the in-house cosmetics. However, we believe that the modeling process can be widely used, as we were able to construct valid ML models for a wide variety of cosmetics. Thus, using formulation data of marketed products, it should be possible to create other models with effective prediction.

The created AI automatically executes the ML models by simply entering formulation information. Although multiple ML models need to be run, it takes only a few minutes to output the prediction results. In fact, some formulators who used the AI said that they were able to generate multiple prototypes in a short period of time and were able to design their desired cosmetics quickly. Especially for less-experienced formulators, the AI was useful to know the important ingredients for controlling tactile sensation intensity. On the other hand, experienced formulators suggested that more prediction accuracy for detailed tactile sensation design would be helpful. In addition, it would be helpful to predict other cosmetics qualities, such as stability and safety. Further investigation based on these opinions is currently being conducted for improved assistance of formulators by the AI.

Conclusion

In this study, we proposed a training cycle with an AI for formulators to efficiently acquire the necessary skills for formulation design. We built the AI contained the ML models that predict physical properties and tactile sensation from formulation data. In order to accurately predict the physical properties and tactile sensations from the formulation information, it was important to analyze data by distinguishing between aqueous formulations and emulsions, to extract essential features by analyzing a large number of features, and to increase the amount of pseudo data. Especially for the tactile sensation, it was important to include predicted values as features in order to take into account the influence of tactile sensations on each other. Our research also reveals that the AI is helpful for formulators to expand their abilities through numerous hypothesis construction and verification. Information provided by the AI is also useful to inspire them. The AI is currently being used for developing a wide range of products and is in the process of being expanded to further improve accuracy and predict other cosmetic qualities by adding data. In the future, we aim at widening the application of the AI to other aspects of cosmetics development as well, such as in the DIY cosmetics setting to assist customers in designing their own products. By expanding the limits of human skills with AI, cosmetic chemists should be able to develop innovative and revolutionary formulations and give cosmetics possibilities of new functions and uses.

Conflict of Interest Statement

This study was funded by POLA Chemical Industries.

References

1. European Commission (2019) The Chemicals Strategy for Sustainability. European Green Deal.
2. Institute of personal care science (n.d.) Certificate in Advanced Cosmetic Science. Retrieved from https://personalcarescience.com.au/CosmeticScience/CertificateinAdvancedCosmeticScience-478/
3. Akiyama Y, et al (2010) Fundamental Study on Quantification of Tactile Sensation. IEEJ Transactions on Electronics, Information and Systems 132: 166-172.
4. Kuilenburg J, et al (2012) An experimental study on the relation between surface texture and tactile friction. Tribology International 48: 15-21.
5. Gilbert L, et al (2013) Predicting sensory texture properties of cosmetic emulsions by physical measurements. Chemometrics and Intelligent Laboratory Systems 124: 21-31.
6. Kikegawa K, et al (2019) Physical origin of a complicated tactile sensation: 'shittori feel'. Royal Society Open Science 6:19039.
7. Varela P, Ares G (2012) Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. Food Research International 148: 893-908.
8. Ares G, et al (2015) Comparison of sensory product profiles generated by trained assessors and consumers using CATA questions: Four case studies with complex and/or similar samples. Food Quality and Preference 45: 75-86.
9. Pemmaraju S, Skiena S (2003) Computational Discrete Mathematics: Combinatorics and Graph Theory in Mathematica. Cambridge, England: Cambridge University Press 336-337.
10. Sikonja R M, Kononenko I (2003) Theoretical and Empirical Analysis ReliefF and RReliefF. Machine Learning Journal 53: 23-69.
11. Xu L, Veeramachaneni K (2018) Synthesizing Tabular Data using Generative Adversarial Networks. arXiv 1811. 11264.
12. Chawla V N, et al (2002) SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research 16: 321-351.

13. Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 785-794.
14. Read J, et al (2009) Classifier Chains for Multi-label Classification. Joint European Conference on Machine Learning and Knowledge Discovery in Databases 254-269.
15. Komatsu Y, Ishimaru S (2015) The Evaluation for Application Sensations of Skin Care Cosmetics: A Method for Objective Evaluation of "Mocchirikan". Journal of Society of Cosmetic Chemists of Japan 49: 309-318.
16. Hayakawa T, et al (2010) Classification Method of Tactile Textures Using Onomatopoeias. Transactions of the Virtual Reality Society of Japan 15: 487-490.