

Design of Petroleum Physical Properties Prediction Application

Harry Budiharjo Sulistyarso¹, Dyah Ayu Irawati², Joko Pamungkas³, Indah Widiyaningsih⁴

^{1,3,4} Department of Petroleum Engineering, UPN "Veteran" Yogyakarta, Indonesia

² Department of Informatics, UPN "Veteran" Yogyakarta, Indonesia

Abstract

Based on the results of previous studies regarding the modeling of the physical properties of petroleum, a mathematical model has been built to calculate the prediction of the physical properties of petroleum. The prediction is based on viscosity, interfacial tension, and density data from the EOR laboratory in UPN Veteran Yogyakarta. The model still cannot be used independently without the Python environment, so to be used easily by more users, the model must be built into an independent application that can be installed on the user's device. In this research, the application design for the physical properties of petroleum prediction application will be carried out. The application is built using the Multivariate Polynomial Regression method according to the model to predict the physical properties of petroleum, and uses Naïve Bayes to classify the petroleum, and *will be the changing result of the physical properties of petroleum modeling that has been made in a previous study*. The shift from model to the application requires several adjustments, including user interface, system, and database adjustments which are implemented as the designs of application. The design is done before the application is built to suit user needs as the result of the research.

Keywords: Prediction Application Design, Physical Properties of Petroleum, Eor, Multivariate Polynomial Regression, Naïve Bayes



This is an open access article under the CC-BY-NC license

INTRODUCTION

In previous study, machine learning-based modeling has been successfully built to measure the physical properties of petroleum from oil sample datasets obtained from the EOR laboratory in UPN Veteran Yogyakarta. In this modeling, two methods are used, namely Multivariate Polynomial Regression and Naïve Bayes Classifier. Multivariate Polynomial Regression is used because the data being modeled is multidimensional and non-linear, such as in calculating viscosity predictions, involving temperature data and biosurfactant data used, and in calculating IFT and density. Meanwhile, the Naïve Bayes Classifier is used to classify oil data into three categories, i.e., light oil, medium oil, and heavy oil (Sulistyarso HB et. al, 2020).

It is becoming increasingly critical in today's quickly changing technology environment that web and desktop apps remain accessible to everyone. The requirements of the application are frequently complex and difficult to comprehend, making it hard for most developers to comply with them (Mueller John, 2003).

The modeling carried out in Python is still a prototype that has not been equipped with a database and can only be used on computers equipped with a Python environment. At the same time, this application needs to be distributed to user's computers to install applications and use them on their computers independently.

In this research, the application design will be built based on the existing model. The application design are database design, system design, and user interface design. Database design is carried out to determine the content and data settings needed to support the system design. System design is carried out to regulate how the system works according to user needs. User interface design is carried out to adjust the appearance to match the predetermined system design.

Corresponding author:

harry_hb@upnyk.ac.id; dyah.ayu.irawati@upnyk.ac.id; joko.pamungkas@upnyk.ac.id

DOI: 10.31098/cset.v1i1.437

Research Synergy Foundation

LITERATURE REVIEW

Several studies referred to in this study include research conducted previously by Sulistyarso et al. (2020), which stated that Naïve Bayes could classify and predict the type of oil produced in the EOR process. In this study, a model has been built to predict the values of the physical properties of petroleum and classify and predict oil categories into 3, namely light, medium and heavy oil. The modeling needs to be further developed into a user-friendly application to be easily understood and used for field purposes by users.

The next research is the research by Roberts and Haddad (2005). Making applications for desktop and mobile has its challenges for user interfaces because, in addition to system development constraints, there are also costs that must be incurred to build a user-friendly interface. The problems encountered and lessons learned during application development are described in the study. Application software architecture, hardware, database, and connectivity are also described.

The following research is research conducted by Obermeyer et al. (1997). In this study, the ALRXSYS expert system application was ported to the active Venus database environment. In the expert system, to integrate the rules and the database, there are interdependencies. The experience of Obermeyer and his team in porting programs to an active database environment shows that code changes must be made to revise the application to operate in a dynamic environment. Code changes made at that time were considered in the moderate category.

The next research is research conducted by Tisdall et al (2018). In this study, it was explained that several things become challenging when porting enterprise applications into applications that can be accessed by mobile devices. These things include how to define the scope of the application to be built, various levels of user experience in using the software, determine the skill level of the developer to be used.

These studies conclude that several changes must be made when porting from modeling to applications. The main difference is managing dependencies between data in the database and aligning them with the user interface.

RESEARCH METHOD

This research was conducted to find a solution to the problem when porting from the physical properties of petroleum prediction model that had been carried out in previous study into an independent application that can meet user needs, be distributed and used efficiently by users. The physical properties of crude oil that will be discussed are viscosity, interfacial tension (IFT), and density.

The steps are taken in this research start from planning the required data sources, data collection, data analysis, and design.

Data Resource

The data were obtained from experiments conducted at the EOR laboratory of UPN Veteran Yogyakarta using crude oil samples from Pertamina's KW-55 well. The physical properties of crude oil obtained from laboratory data are experimental data to get the amount of viscosity, IFT, and density by changing the temperature and concentration of the given biosurfactant. These data are used as raw data for the prediction process. Table 1 shows the viscosity data obtained from laboratory trials and then the data is used as a raw data dataset in this study. Table 2 shows the IFT raw data, and tabel 3 shows density raw data

The raw data used in the viscosity model with 5% biosurfactant in multivariate polynomial regression include:

Table 1. Viscosity Raw Data

No	Biosurfactant	Temp	Viscosity	Oil Category
1.	5 %	30	0.95	Light Oil
2.	5 %	40	0.85	Light Oil
3.	5 %	50	0.75	Light Oil
4.	5 %	60	0.65	Light Oil
5.	5 %	70	0.55	Light Oil
6.	5 %	80	0.25	Light Oil

The raw data used in the IFT model with 5% biosurfactant in multivariate polynomial regression include:

The raw data used in the density model in multivariate polynomial regression include:

Table 2. IFT Raw Data

No	Biosurfactant	Temp	IFT
1.	5 %	30	10.35
2.	5 %	40	8.27
3.	5 %	50	7.6
4.	5 %	60	5.32
5.	5 %	70	4.3
6.	5 %	80	3.19

Table 3. Density Raw Data

No	Biosurfactant	Temp	Density
1.	0 %	60	0.87
2.	15 %	60	0.89
3.	2.5 %	60	0.88
4.	5.0 %	60	0.91
5.	10.0 %	60	0.93

The multivariate polynomial regression formula used is represented in the equation below:

$$y = a_0 + a_1 * x_1 \dots \dots \dots \text{equation 1}$$

Where y is the dependent variable, a0 is the constant, a1 is the coefficient for x1, and x1 is the independent variable. The formula in equation 1 represents a regular regression polynomial.

$$y = a_0 + a_1 * x_1 + a_2 * x_2 + \dots \text{large} \dots + a_n * x_n \dots \dots \dots \text{equation 2}$$

While equation 2 represents multivariate polynomial regression, which has a function similar to ordinary polynomial regression but has n independent variables according to the features of variable x, which is used to predict variable y.

Data Analysis

From the data obtained and the interview with the EOR researchers, an analysis is carried out to determine the user's needs for the software. User requirements are as follows:

1. Can predict on viscosity data
2. Can predict the IFT data
3. Can predict density data
4. Can save prediction results data
5. Can see the distribution of IFT training data
6. Can see the distribution of the viscosity training data
7. Can see the distribution of density training data
8. Can see the distribution of IFT test data
9. Can see the distribution of viscosity test data
10. Can see the distribution of density test data

Design

A system design is made from the needs of existing users, including system flow design, database design, and user interface design. To port from modeling to application, what needs to be done is as follows:

- Do the database system planning process by creating relationships under existing data modeling.
- After the database design is created, the next step is to design the application mockups according to user needs and what actions the user can take.
- After planning the mockup, the next step is to design the data modeling that will be used, such as taking the features used in each viscosity, IFT, and density data model.
- The next step is to implement the relevant data modeling into the Python programming language.
- After modeling the data against the programming language, the next step is to check the validity of the data on each data model that is made, whether it is under actual data or viceversa.
- When the results of the validity of the resulting data do not match, the next step is to return to data modeling to produce data with the best validity for each data model.
- After planning the validity of the data, the next step is to build the application into a desktop form.
- Checking the suitability of the mockup with the application.
- Next, build a desktop application in *.exe and *.dmg to distribute to the user.

FINDING AND DISCUSSION

A. System Design

A flowchart describes the design of the system flow in figure 1. The system will start to test the raw data from the laboratory in array format, then user can choose one option of the menus. After the choice is made, the system will perform calculations using multivariate polynomial regression to get prediction results from the entered data. The prediction result will be classified by the Naïve Bayes in the system into three types of oil, i.e. light oil, medium oil and heavy oil.

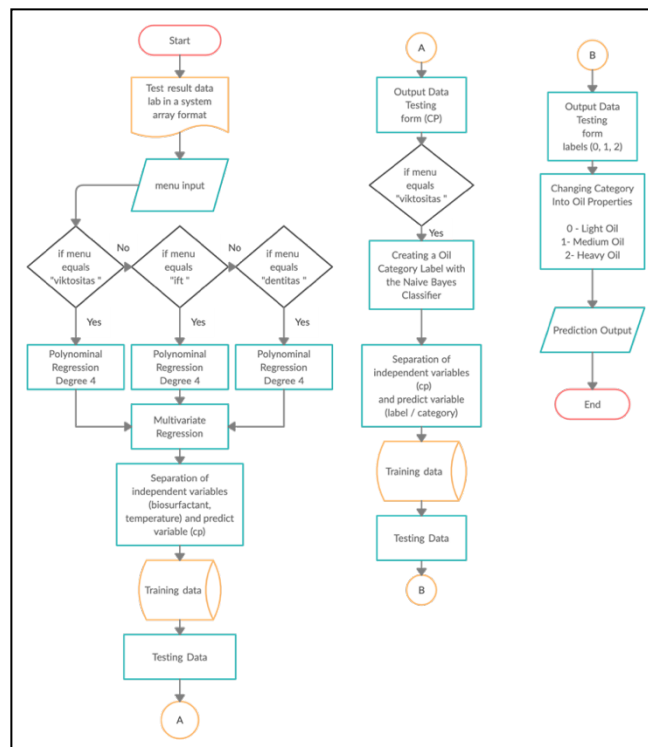


Figure 1. System Flowchart

B. Database Design

The following is a representation of the database design used in the application, where the plan includes:

- Tbl_of_categori functions as a container for the category of petroleum properties according to the value of the existing earth properties.

- Tbl_of_model represents the data model used, such as IFT, Density, or viscosity in the application.
- Tbl_of_data_model represents the data model used, namely Training, Testing, and Validation data on the application, functions as the data category.
- Tbl_of_data represents the results of the predicted data and also the training data in the application.

The relationship in database design below includes:

- For each tbl_of_categori contains one or more data in each tbl_of_data
- For each tbl_of_model has one or more data in each tbl_of_data
- For each tbl_of_data_model has one or more data in each tbl_of_data
- As for multiple data, it must have one tbl_of_model_data, tbl_of_model, tbl_of_categories.

The ERD diagram in Figure 2 illustrates the database design.

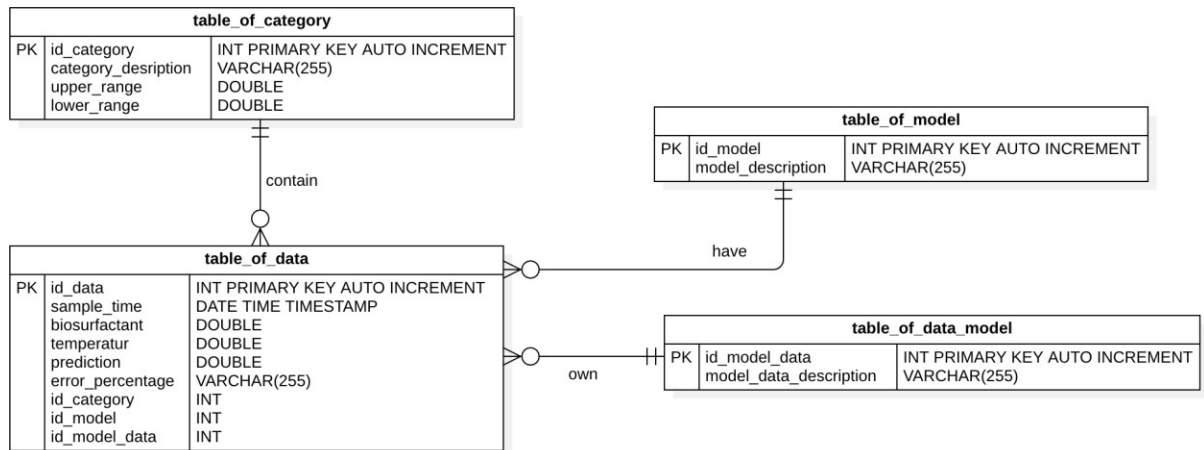


Figure 2. Entity Relationship Diagram

C. User Interface Design

In this section, several mock-up designs for the prediction of petroleum physical properties application are presented.

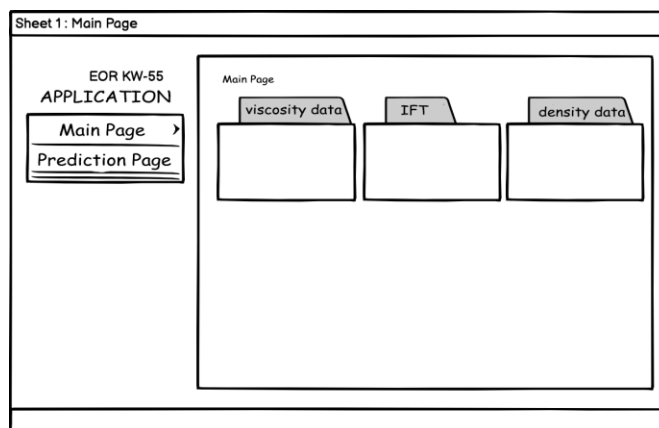


Figure 3. Main Menu Mock-up

Figure 3 represents the main menu display in the application, where this view includes the distribution of viscosity, ift, and density data wrapped in each menu box, and a prediction sub menu to predict data that has been modeled on the application.

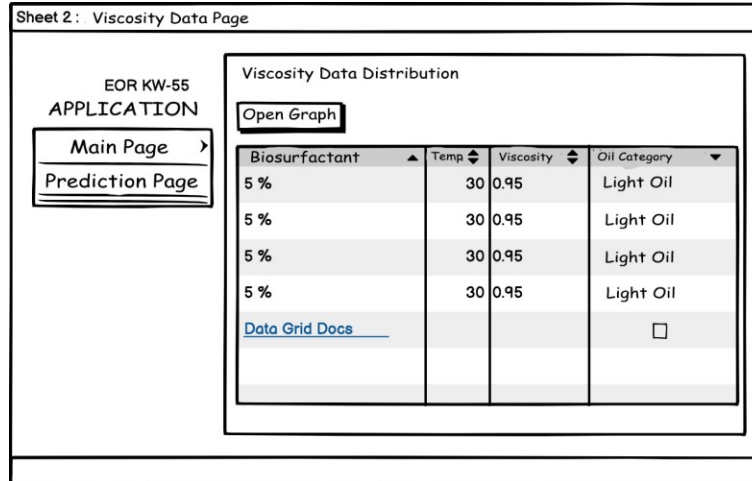


Figure 4. Viscosity Data Distribution Mock-up

Figure 4 displays the distribution of raw / training viscosity data is the distribution of viscosity data owned by the application which will later become a reference in predicting viscosity data.

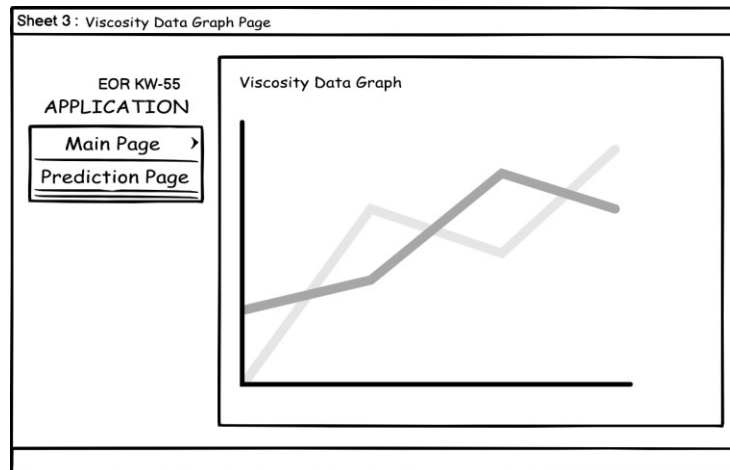


Figure 5. Viscosity Data Distribution Graphic Mock-up

Figure 5 represents graphical display of the distribution of raw / training viscosity data is a display that displays a graph of the distribution of viscosity data on the application.

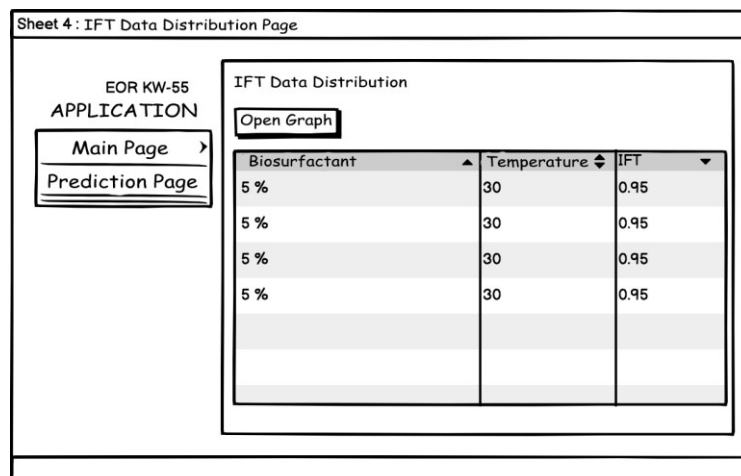


Figure 6. IFT Data Distribution Mock-up

Figure 6 displays the distribution of IFT raw/training data. The distribution of IFT data will later become a reference in predicting IFT data.

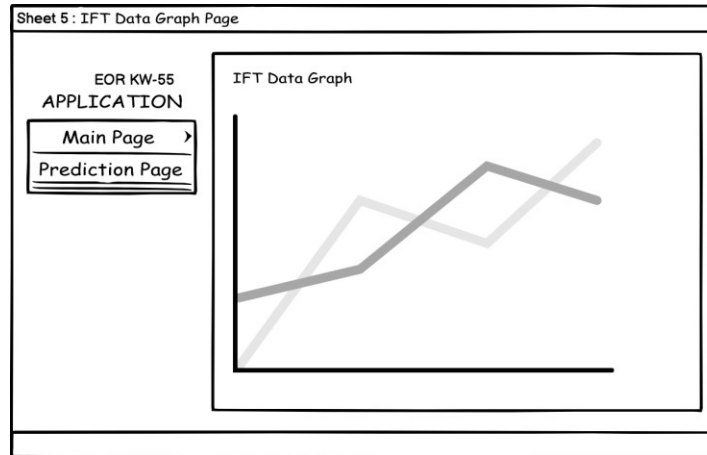


Figure 7. IFT Data Distribution Graphic Mock-up

Figure 7 represents graphical display of the distribution of IFT raw / training data. This is a display that displays a graph of the distribution of IFT data on the application.

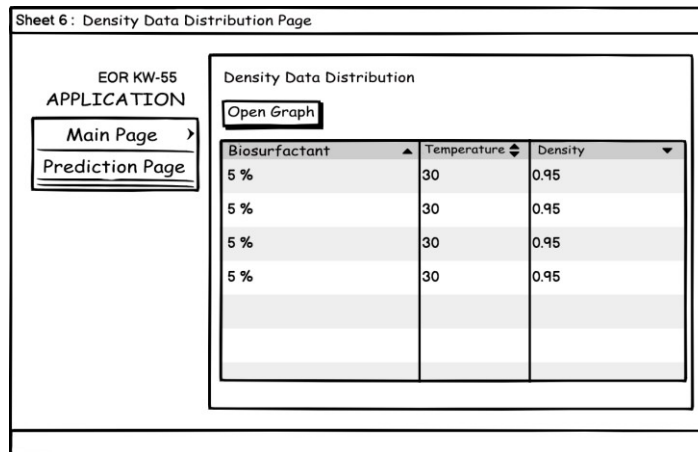


Figure 8. Density Data Distribution Mock-up

Figure 8 displays the distribution of density raw/training data. This is the distribution of density data which will later become a reference to predict viscosity data. Figure 9 displays the graphic of density raw/training data.

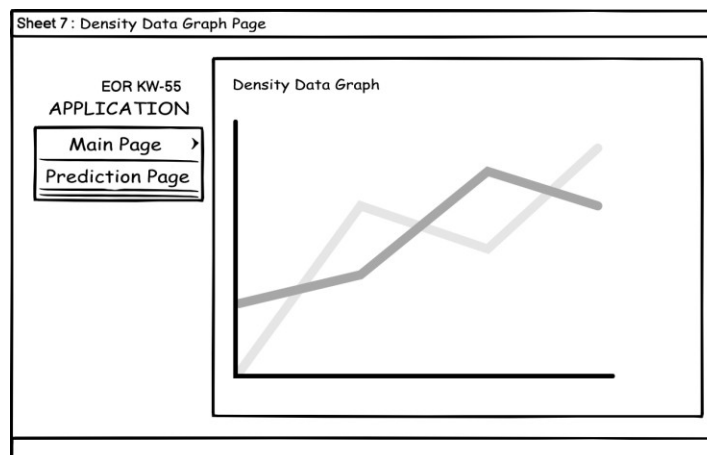


Figure 9. Density Data Distribution Graphic Mock-up

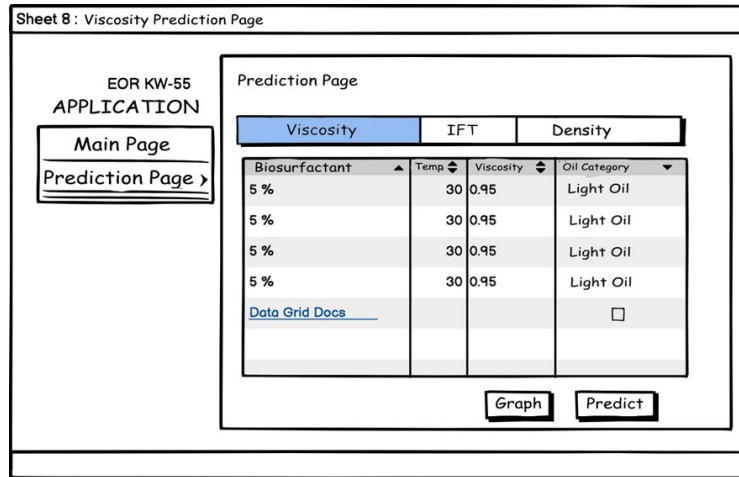


Figure 10. Viscosity Prediction Mock-up

Figure 10 is the the viscosity prediction menu display. This display mock up is used to predict new viscosity data, the results of the viscosity prediction depend on the viscosity training data model used in the application. Figure 11 represents the graph of the viscosity prediction data.

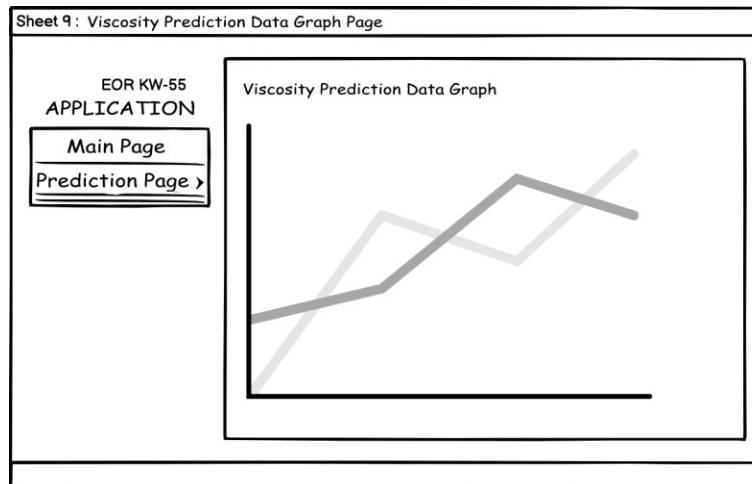


Figure 11. Viscosity Prediction Result Graphic Mock-up

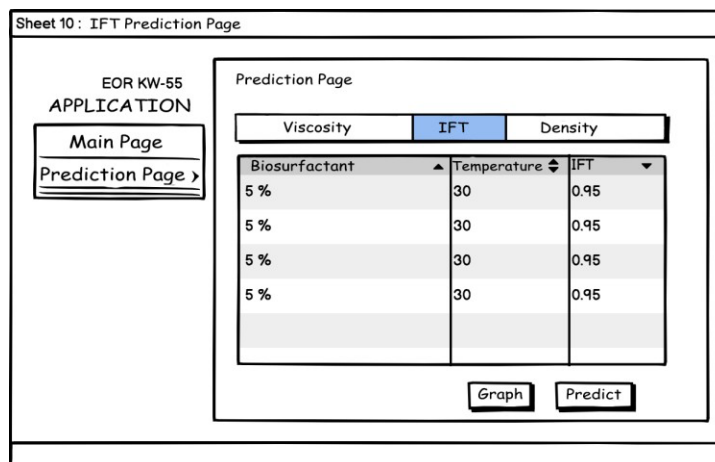


Figure 12. IFT Prediction Mock-up

Figure 12 is the the IFT prediction menu display. This display mock up is used to predict new IFT data. The results of the IFT prediction depend on the IFT training data model used in the application. Figure 13 represents the graph of the IFT prediction data.

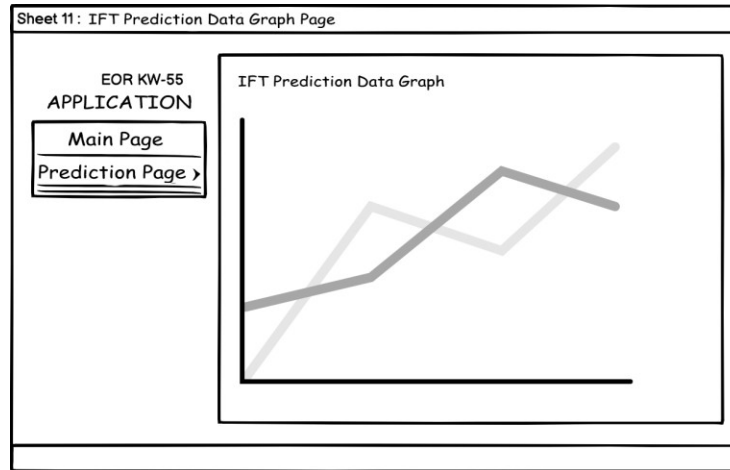


Figure 13. IFT Prediction Result Graphic Mock-up

Figure 14 is the the density prediction menu display. This display mock up is used to predict new density data. The results of the density prediction depend on the density training data model used in the application. Figure 15 represents the graph of the density prediction data.

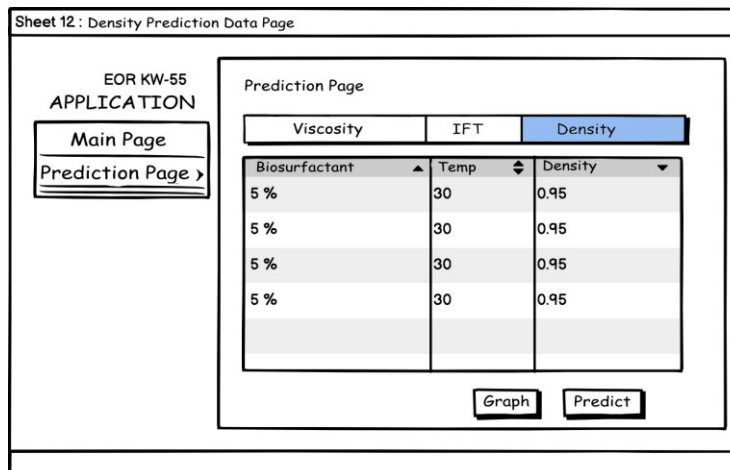


Figure 14. Density Prediction Mock-up

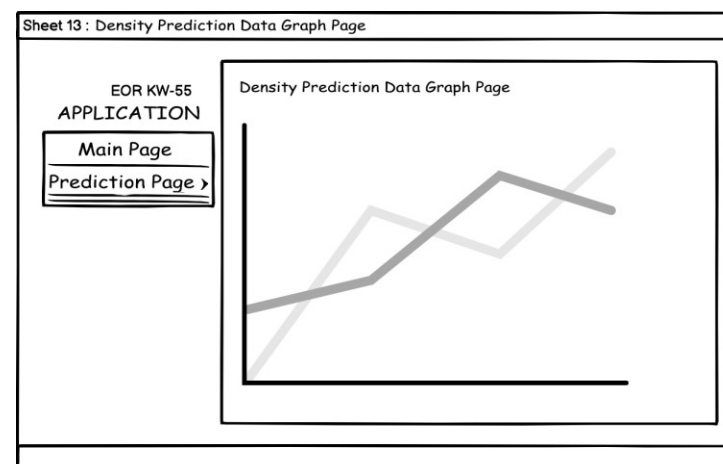


Figure 15. Density Prediction Result Graphic Mock-up

CONCLUSION AND FURTHER RESEARCH

The design of the application for predicting the physical properties of petroleum has been carried out by the data obtained and has been under the user's needs. The application will be built by the design result in this research.

In the future, this research can then be developed into a desktop based applicatopn or mobile-based application using Android, so that researchers in the petroleum sector can use applications in the field.

REFERENCES

- Bustami., (2013), Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi, *TECHSI : Jurnal Penelitian Teknik Informatika*, Vol. 3, No.2, Hal. 127-146.
- Irawati DA, Syaifudin YW, Tomasila FE, Setiawan A, Rohadi E. 2018. "Development of Android-based Rabbit Disease Expert System". *International Journal of Engineering and Technology*. 7(4.44):82-87.
- Kementerian ESDM. 2015. Rencana Strategis 2015-2019. Direktorat Jenderal Minyak dan Gas Bumi Kementerian Energi dan Sumber Daya Mineral Indonesia.
- Mueller John. 2003. Accessibility for Everybody. Apress. ISBN : 978-1-4302-5448-5. Springer:Switzerland.
- Obermeyer L, Warshaw L, dan Miranker Daniel P. 1996. "Porting an Expert Database Application to An Active Database : An Experience Report". CIKM 96 : Proceeding of The Workshop on Databases: active and real time:37-40.
- Robert T dan Haddad HH. 2005. *Application Development for Desktop and PDA-style Device: Issues and Lessons Learned*. Proceeding of The 43th ACM Southeast Conference, Vol 2:242-247. Pages 242-247. <https://doi.org/10.1145/1167253.1167309>.
- Saleh, A. 2015. Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga Citec Journal, Vol. 2, No. 3, ISSN: 2354-5771
- Sarkar S dan Sriram Ram S. 2001. Bayesian Models for Early Warning Bank Failures. *Management Science Journal*. 47(11).
- Satter A, Varnon JE, dan Hoang MT. 1994. Integrated Reservoir Management. *Journal of Petroleum Technology*. 46(12):1057-1064.
- Sen, R. 2008. "Biotechnology in petroleum recovery:The Microbial EOR". *Progress in Energy and Combustion Science Journal*. 34(6):714-724.
- Sulistyarso HB, Irawati DA, Pamungkas J, Widiyaningsih I. 2020. Modeling of Crude Oil Types Classification Using the Naive Bayes Classifier Method. *Proceeding of LPPM UPN "Veteran" Yogyakarta Conference Series 2020-Engineering and Science Series*: 328-339. LPPM UPNVY. Yogyakarta.
- Sulistyarso HB, Pamungkas J, Gusmarwani SR, Wahyuningsih T. 2019. Aplikasi Biosurfaktan dalam Upaya Peningkatan Perolehan Minyak Tahap Lanjut : Uji Laboratorium pada Sampel Sumur KW-58. *Prosiding SNCPP 2019 dengan Tema Pengembangan Ristek dan Pengabdian menuju Hilirisasi Industri*: 55-60. LPPM UPNVY. Yogyakarta
- Tisdall T, Chobharkar P, dan Kim D. 2018. Get Mobile: GetMobile: Mobile Computing and Communications. <https://doi.org/10.1145/3229316.3229324>.
- Vural MS dan Gök M. 2017. Criminal Prediction using Naïve Bayes Theory. *ACM Journal Neural Computing and Applications*. 28(9).