Proceedings of the
International Conference on Mechanical Engineering 2009
(ICME2009) 26- 28 December 2009, Dhaka, Bangladesh

ICME09-FM-02

# PREDICTION OF STANDPIPE PRESSURE USING REAL TIME DATA

**Dipankar Chowdhury[1], Pål Skalle[1]** and **Mahbubur Rahman[2]**

[1]Department of Petroleum Engineering and Applied Geophysics,
Norwegian University of Science & Technology, NO- 7491, Trondheim, Norway
[2]Department of Petroleum and Mineral Resources Engineering,
Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

## ABSTRACT
Standpipe pressure (SPP) represents the total frictional pressure drop in the hydraulic circuit used for rotary drilling operation. Conventional approach for SPP calculation is based on a number of simplifying assumptions that cannot be realized in real drilling operation. Hence the conventional approach does not produce SPP estimates with sufficient accuracy. The limitations of conventional approach led to the development of alternative methods for SPP estimation using real time data. In this paper, SPP was estimated by regression models and by instance-based reasoning (IBR) models using real time data collected while drilling a deviated well in the North Sea. These models were developed using mud flow rate and bit measured depth as independent parameters. For the real time data considered in this paper, the IBR models produced much better SPP estimates than the regression models.

**Keywords**: Standpipe pressure, Regression, Instance-based reasoning, Real time drilling data

## 1. INTRODUCTION
Rotary drilling is the most widely used drilling method. Drilling fluid is an essential component of a rotary drilling system. It is circulated through different parts of the hydraulic circuit by the mud pump principally for removing cuttings. During its circulation, pressure drop occurs at different segments of the hydraulic circuit due to friction between the drilling fluid and the surface in contact. The total frictional pressure drop in the hydraulic circuit, which equals the pump pressure, is termed as standpipe pressure (SPP). The hydraulic circuit typically consists of standpipe, rotary hose, swivel, kelly, drillstring, drill bit, and annular section between the drillstring and the casing or borehole wall.

SPP is an important drilling parameter that must be known with sufficient accuracy for selecting proper jet bit nozzle size, determining optimum flow rate to ensure efficient hole cleaning and for selecting proper mud pump liner. Continuous monitoring of SPP also helps in identifying downhole problems. For example, too low SPP can be caused by washed out pipe or bit nozzle, loose joint or broken drill string, worn pump packing or liner, or lost returns due to formation fracture. On the other hand, too high SPP could indicate a plugged drill bit or increased mud density or viscosity. SPP anomalies can provide an early warning of circulation problems and thus can help the driller to make corrections before the situation is out of control.

SPP can be estimated using analytical frictional pressure drop relations. These relations are available for the three widely used rheological models namely Bingham plastic, Power law and Herschel-Bulkley. A detail description of this conventional approach is presented in [1]. However, the relations used in the conventional approach are based on a number of simplifying assumptions, such as concentric annular and circular sections, non-rotating drillstring, isothermal conditions in the bore hole and steady state axial flow. These simplifying assumptions are not completely valid in real life [2]. The effect of pipe eccentricity, pipe rotation, and temperature and pressure variations can have significant effect on frictional pressure drop in the annulus. Accurate determination of SPP is necessary for safe drilling because too high SPP can fracture the formation resulting in lost circulation and too low SPP can cause kick which can lead to a blow out.

The limitation of the conventional approach led to the development of alternative method of SPP prediction using real time data. In this paper, the methodology used for estimating SPP using real time drilling data for a deviated well drilled in the North Sea is presented. Two approaches are considered for this purpose – i) regression analysis and ii) instance-based reasoning (IBR).

## 2. REAL TIME DATA COLLECTION

Real time data are collected by measurement while drilling (MWD) systems. Unlike wireline logging, MWD systems can be used for gathering and transmitting data from bottom-hole back to the surface without any interruption of the drilling operation. The data transmitted can be directional data, data related to the petrophysical properties of the formations and data related to drilling parameters such as SPP, weight on bit (WOB), mud flow rate, downhole temperature and torque. Relevant sensors and transmission equipment housed in a non-magnetic drill collar in the bottom hole assembly are used for data collection and transmission to the surface. A typical MWD system consists of a downhole system (power source, sensors, transmitter and control system), a telemetry channel (mud column used for sending pulses to surface) and a surface system (used for detecting pulses, decoding the signal and presenting numerical display, geological log etc.). A typical MWD system is shown in Fig 1.

The real time drilling operation data used in this paper were collected while drilling the 8.5¨diameter deviated section of a well drilled in the Gullfaks field of the North Sea. The section was drilled using an oil based mud. Twelve measurements for each of the 37 drilling parameters (including SPP, WOB, bit measured depth, bit torque, travelling block position, mud flow rate, mud density, equivalent circulating density, hook load and gas content in the mud) were made simultaneously every minute. A total of 386,128 measurements were made for each drilling parameter from a measured depth of 5060 m to 6221 m. The collected data were analyzed using MATLAB.

### 2.1 Parameter Selection and Data Segregation

Among the different drilling parameters measured, mud flow rate and bit depth (DBTM) were considered for developing regression and IBR models. These are two of the four parameters that principally affect the pressure drop in a wellbore. The other two parameters are density of the fluid and flow

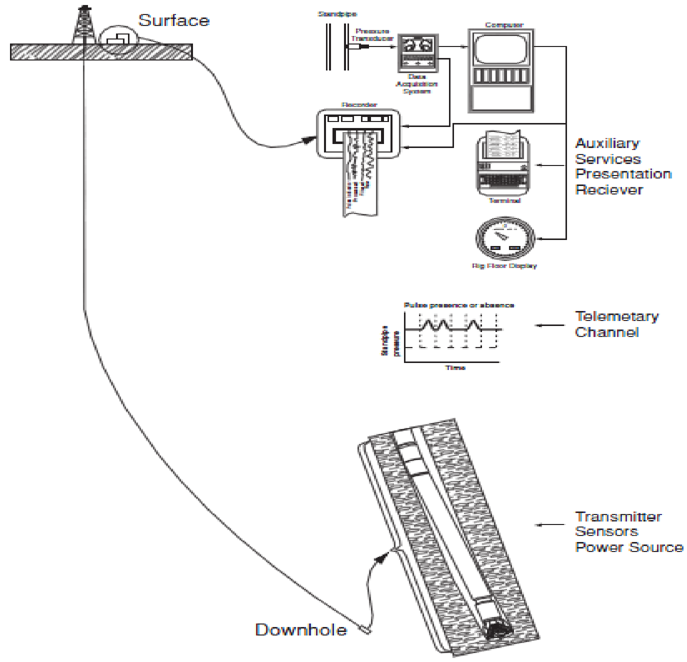



Fig 1. A typical MWD system[3]

area. Since an oil based mud was used throughout the operation, and the dimensions of the drillstring and the hole remained the same during data collection, the last two parameters were not considered for model development.

The collected data were divided into two groups: data collected during tripping-in and data collected during drilling. This was necessary because tripping-in (i.e. the process by which the drillstring is run into the hole) and drilling (i.e. the process by which the formation is penetrated by the drill bit) represent two different phenomena. During tripping-in, mud is circulated at times at constant flow rates with or without rotating the drillstring keeping the bit off bottom to clean up the hole. The flow can be either laminar or turbulent. On the other hand, dynamic condition prevails in the hole during drilling as the drillstring is rotating and vibrating. Hence the flow is usually turbulent during drilling.

Measured depth (i.e. the distance measured along the path of the wellbore) and bit measured depth (i.e. the distance travelled by the drill bit along the path of the wellbore) were used for data segregation. For tripping-in data, the two depths were considered different. But they were assumed to have an absolute difference of less than or equal to 0.1 m for the data collected during drilling. Furthermore, WOB (i.e. the load put on the bit by the drill collars to make penetration) was zero for tripping-in data and nonzero for drilling data. Also, the bit torque assumed nonzero values for drilling data while it was either zero or nonzero for tripping-in data.

### 2.2 Determination of Stationary SPP

To detect stationary SPP, determination of stationary flow rate and travelling block position is necessary. A flow rate was considered stationary if it had an absolute difference of less than or equal to 0.1 lpm (1 lpm = 0.000167 $m^3/s$) with its previous or next measured flow rate. The block position was considered stationary if it had an absolute difference of less than or equal to 0.1 m with its previous and next measured block positions. The SPP was considered stationary if the absolute difference between the previous measured value and the present one, and that between the present measured value and the next one was less than or equal to 0.1 bar.

Using the criteria mentioned earlier, 10035 measured mud flow rate, DBTM and SPP during tripping-in, and 4320 measured mud flow rate, DBTM and SPP during drilling were found. However, several similar measured data sets were found at particular bit depths. This happened because the measurements were made after a short time interval of five seconds. To overcome this, the mean of all the flow rates at a particular bit depth interval of one meter was calculated and only one flow rate closest to the mean was determined. The SPP corresponding to this flow rate was considered for further work. This action reduced the number of tripping-in data set to 828. The same refinement when used for drilling data reduced the number of data set to 729. After removal of outliers (i.e. observations that are indicative of measurement error), the number of data set was further reduced to 764 for data collected during tripping-in and to 728 for data collected during drilling.

### 2.3 Selection of Training Examples and Query Instances

Training examples constituted of measured mud flow rate, DBTM and SPP were used for model development. Query instances consisting of the same parameters mentioned earlier were used for SPP prediction using the developed regression and IBR models. Among the 764 tripping-in data sets, every fifth data set is selected as a query instance including the first data set. This resulted in 611 training examples and 153 query instances. Same procedure produced 582 training examples and 146 query instances for the data collected during drilling.

## 3. MODEL DEVELOPMENT

To estimate SPP for the tripping-in query instances, four regression models and nine IBR models were used. Same approach was used for the data collected during drilling.

### 3.1 Regression Analysis

Regression analysis is a statistical technique for investigating functional relationship between the response variable, and one or more predictor variables. Regression models are most commonly developed using an error minimization criterion known as *Least Squares Regression*. In this method, the regression models are produced by minimizing the sum of the squares of the residuals. These models are based on three principal assumptions - independent predictor variables, equally reliable observations and normally distributed residuals. There are a number of least squares regression methods, such as linear least squares, weighted least squares, robust least squares and nonlinear least squares. Linear and nonlinear least squares methods produce regression models with acceptable goodness of fit parameters (such as SSE, RMSE and R-square) when none of the assumptions mentioned earlier is violated.

The parameters of a nonlinear model are adjusted using a fitting algorithm. Most widely used fitting algorithm is *Trust-Region*. This algorithm must be used if coefficient constraints are present. Another popular algorithm is *Levenberg-Marquardt*. It can be tried when the Trust-Region method does not produce a reasonable fit and there is no coefficient constraint [4].

There are several robust regression schemes. But the most widely used schemes are *Least Absolute Residuals (LAR)* and *Bisquare Weights*. In LAR scheme, a curve is found that minimizes the absolute value of the residuals instead of the squared differences. Bisquare weights scheme, on the other hand, minimizes a weighted sum of squares.

For a regression model that fits the observed data well, SSE is close to zero, R-square is close to one and RMSE is close to zero.

#### 3.1.1 Regression Models for Tripping-in Data

Following four regression models were developed-

Model1: $SPP = 0.0001521Q^{1.873} + 18.44$

Model2: $SPP = 0.002844Q^{1.497}$

Model3: $SPP = 20.2 + 2.46\times10^{-8}(DBTM)(Q^{1.75}) + 0.000241Q^{1.75}$

Model4: $SPP = 15.2 + 8.42\times10^{-10}(DBTM)(Q^{1.76}) + 0.000358Q^{1.76}$

Both Model1 and Model2 were developed using robust least squares regression with LAR scheme and Trust-region algorithm. The curve fitting toolbox of MATLAB was used for this purpose. Model3 and Model4 were multiple regression models developed using linear least squares regression and robust least squares regression respectively. All the models were developed using MATLAB

Among the four models, Model1 fitted the training data well with SSE=3488, R-square=0.9992 and RMSE=2.395. On the other hand, Model4 produced the worst fit of the training data with SSE=195000, R-square=0.955 and RMSE=17.9.

### 3.1.2 Regression Models for Drilling Data

Similar procedure as tripping-in data was followed for drilling data. Four regression models were developed as before. They are designated as ‘Model1D’, ‘Model2D’, ‘Model3D’ and ‘Model4D’. The four models are as follows-

Model1D: $SPP = 0.0002494Q^{1.811} + 10.31$

Model2D: $SPP = 0.0005629Q^{1.709}$

Model3D: $SPP = -162 - 1.57\times10^{-6}(DBTM)(Q) + 0.216Q$

Model4D: $SPP = -168 + 1.32\times10^{-6}(DBTM)(Q) + 0.202Q$

Model1D and Model2D were found using nonlinear least squares regression. Model3D was found using linear least squares regression while Model4D was found using robust least squares regression. Among the four models, Model1D fitted the training data well with SSE=1213, R-square=0.9864 and RMSE=5.578. On the other hand, Model4D produced the worst fit of the training data with SSE=37500, R-square=0.924 and RMSE=8.05.

## 3.2 Instance-based Reasoning

Instance-based reasoning (IBR) is a machine learning method. Machine learning aims at developing computer programs which will learn from an existing database formed by real life experience and predict the value of the desired parameter (i.e. output) for a new unobserved situation using its learning experience.

The following methods were used to develop IBR models to estimate SPP for the deviated long well drilled in the North Sea:

- k-nearest neighbor learning algorithm
- Distance-weighted k-nearest neighbor learning algorithm
- Locally weighted linear regression

The k-nearest neighbor learning algorithm is the most basic IBR method [5]. It predicts the output parameter based on the k-nearest neighbors of the query instance. The value of k can be any positive integer starting from one to N. When k=N, the prediction becomes the global average of all the training examples. The most frequently used value of k is three.

Distance-weighted k-nearest neighbor learning algorithm is a refinement to the previous learning algorithm. In this method, the contribution of each of the k-nearest neighbors is weighted with respect to its Euclidian distance from the query instance.

In locally weighted linear regression method, a different local approximation is calculated for each distinct query instance. In this method, weights are determined by minimizing the error in prediction using the training examples near the query instance. For error minimization, Quasi-Newton method and genetic algorithm were used in the current work. The Quasi-Newton method uses gradient to search for the minimum point of the objective function while genetic algorithm uses parallel search technique to find the optimal point.

A detail description of all these IBR methods can be found in [3] and [5].

Nine IBR models for tripping-in data and nine IBR models for drilling data were developed. The basic structure of these models is shown in Fig 2. Every model worked on the basis of an instance base. It was a database containing both training examples and the query instances. The instance base was formed with the help of MS Excel. The Euclidean distance between the query instance and a training example was determined using two attributes (DBTM and mud flow rate) for every query instance. The Euclidean distance was used to find the nearest neighbors and also to calculate the weights for every training example for distance-weighted models. For locally weighted models, the attributes were weighted.

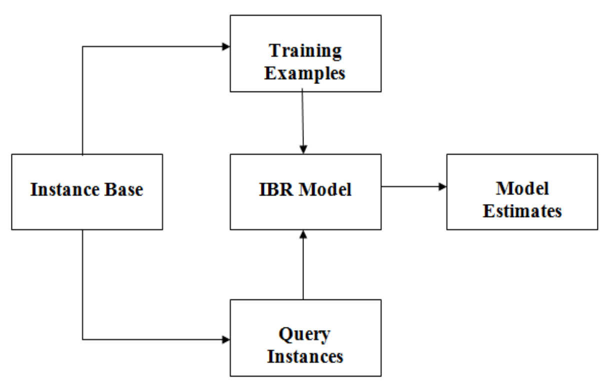



Fig 2. Basic structure of IBR models

Among the nine IBR models developed for each of the two groups of data, four models were based on the k-Nearest neighbor learning algorithm, three were based on distance-weighted learning algorithm and two were based on locally weighted linear regression.

The locally weighted linear regression model was of the following form:

$$SPP = w_o + w_1 DBTM + w_2 Q \qquad (1)$$

The weights were found by minimizing the following distance weighted error function for the three nearest neighbors of the query instance in the Euclidean space:

$$Error = \sum_{i=1}^{3} \frac{(SPP_{measuredi} - SPP_{estimatei})^2}{d(x_q, x_i)} \qquad (2)$$

Here $SPP_{measuredi}$ is the measured SPP and $SPP_{estimatei}$ is the estimated SPP for the ith nearest neighbor. For error minimization, two MATLAB functions fminunc() and ga() were used. The function fminunc() uses Quasi-Newton algorithm while ga() uses genetic algorithm for minimizing the error function. The weights corresponding to minimum error were used in Eq. (1) for estimating SPP for the query instance.

A detail of the IBR models used can be found in [3].

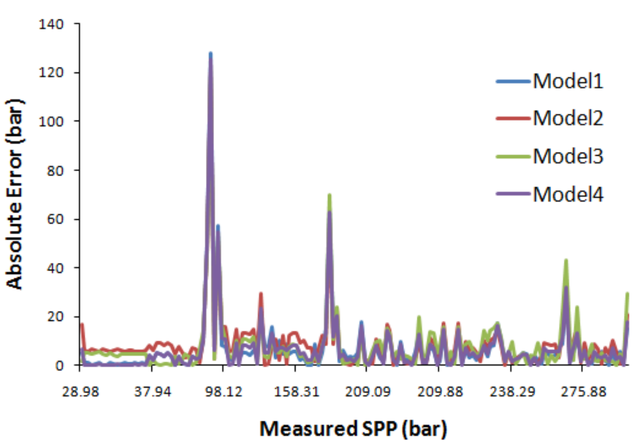
## 4. RESULTS AND DISCUSSION

The results obtained for tripping-in data and drilling data are presented in this section's two subsections.

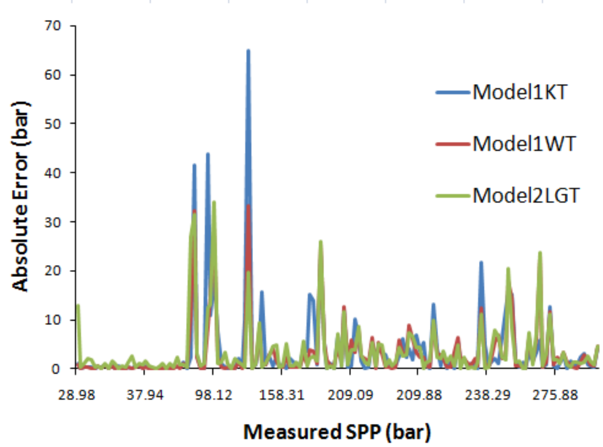
### 4.1 Tripping-in Data

The absolute error in predicted SPP for the 153 tripping-in query instances is presented in Fig 3. Among the nine IBR models, Model1KT based on k-nearest neighbor learning algorithm outperformed the other three k-nearest neighbor models. Similarly, Model1WT among the distance-weighted models and Model2LGT among the linear regression models provided good estimates of SPP. For Model1KT, the data were not normalized (i.e. dividing all the values of an attribute by the largest value among them). The model also did not include updating the training example database (i.e. addition of the current query instance to the training example database). Model1WT estimated SPP for the query instance based on the weighted average of the three nearest neighbors of the query instance. Model2LGT predicted SPP using Eq. 1. The optimum weights were found by genetic algorithm using the three nearest neighbors of the query instance. Model1WT and Model2LGT used normalized data. They also updated the training example database after estimating the SPP for a query instance.

Fig 3 shows that the regression models predicted SPP with absolute errors in the range of zero to 20 bars while the IBR models predicted SPP with absolute errors in the range of zero to 10 bars for most of the query instances. A close observation of the error plots reveals that the regression models produced SPP estimates with errors greater than three bars for a large number of instances. It is 92 for Modle1, 116 for Model2, 108 for Model3 and 91 for Model4. On the other hand, Model1KT produced SPP estimates with errors greater than three bars for 42 query instances, Model1WT did it for 41 instances and Model2LGT did the same for 45 instances. These numbers are approximately half of those observed for the four regression models.

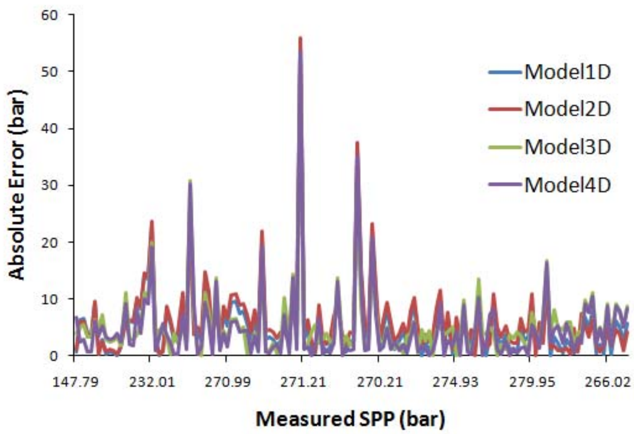



(a)

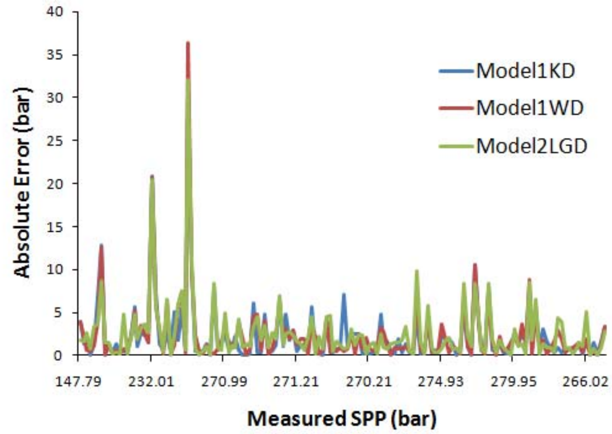



(b)

Fig 3. Comparison of absolute error for (a) tripping-in regression models and (b) tripping-in IBR models



(a)



(b)

Fig 4. Comparison of absolute error for (a) drilling regression models and (b) drilling IBR models

### 4.2 Drilling Data

Similar results were obtained for the drilling data. The results are shown in Fig 4. Fig 4 shows that the highest absolute error in predicted SPP was well over 50 bars for regression models while for IBR models it was around 35 bars. A close observation of the two error plots reveals that Model1D produced SPP estimates with absolute error greater than three bars for 80 query instances, Model2D did it for 101 instances, Model3D did the same for 92 instances and Model4D did it for 68 query instances. On the other hand, Model1KD produced SPP estimates with absolute error greater than three bars for 30 instances, for Model1WD it was 33 and for Model2LGD this number was 39. All these numbers are even less than half of those for the regression models. This indicates the credibility of IBR models as better estimators of SPP than the regression models for the present data.

## 5. CONCLUSION AND FURTHER RECOMMENDATIONS

Due to the limitations of the conventional approach, alternative methods for SPP estimation using real time data were developed. IBR is one such method. In this paper, IBR models along with regression models are presented. For the real time data used in this paper, IBR models produced more accurate estimates of SPP compared to the regression models tried.

However, the results obtained are limited by the fact that the data used were collected while drilling the deviated segment of a well using an oil based mud. Therefore, it is recommended that the regression analysis and IBR approach should be tried with real time data collected while drilling wells in different geological locations using different drilling muds. Also, other machine learning methods such as case-based reasoning (CBR) and neural network can be tried. Furthermore, the results obtained by the IBR approach can be compared with those obtained by the CBR and neural network approach.

## 6. ACKNOWLEDGEMENT

A special thanks to Dr. Sigve Hovda of Verdande Technology AS for his support during the work.

## 7. REFERENCES

1. Chowdhury, D., 2008, ¨Hydraulic Pressure Losses in Drilling Fluid¨, MSc Project Report, Norwegian University of Science and Technology, Trondheim, Norway.
2. Bourgoyne, Jr. A. T., Chenevert, M. E., Millheim, K. K. and Young, Jr. F.S., 2003, *Applied Drilling Engineering*, Society of Petroleum Engineers, Texas, USA, pp. 12-17,131-157.
3. Chowdhury, D., 2009, ¨Prediction of Standpipe Pressure Using Real Time Data¨, MSc Thesis Report, Norwegian University of Science and Technology, Trondheim, Norway.
4. *Curve Fitting Toolbox User's Guide* (Version 1), The MathWorks Inc., 2002, Massachusetts**,** USA, pp 1-40 (Chap. 3).
5. Mitchell M. T., 1997, *Machine Learning*, The McGraw-Hill Companies Inc., Singapore, pp. 1-18, 20-45, 230-245.

## 8. NOMENCLATURE

| Symbol | Meaning | Unit |
|---|---|---|
| Q | Mud flow rate | ($m^3$/s) |
| DBTM | Bit measured depth | (m) |
| N | Number of training examples | dimensionless |
| $w_o$, $w_1$, $w_2$ | Weights | dimensionless |
| $d(x_q, x_i)$ | Euclidean distance between normalized query instance and training example | dimensionless |