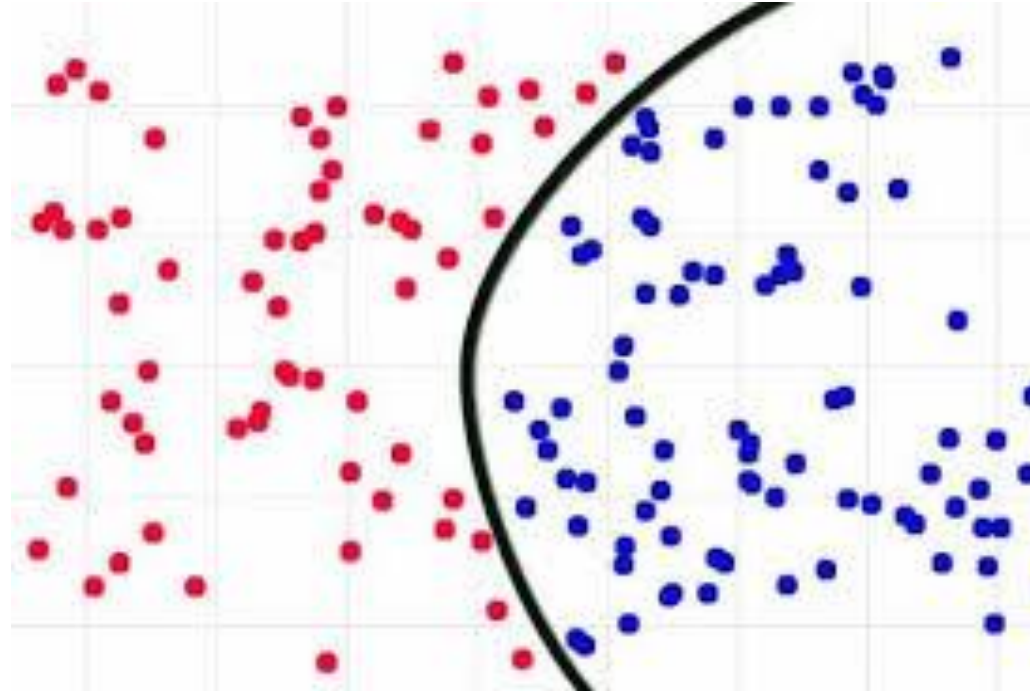
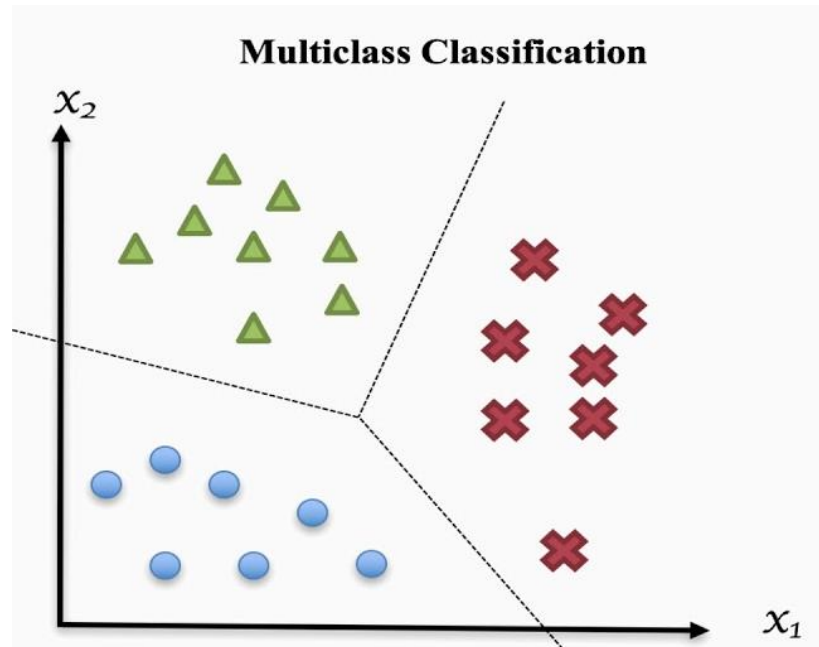


# Classification, Regression and Clustering

# Introduction

- When a machine learning task is presented, the first thing is to know whether the learning task is Classification or regression or clustering problem so that next you can pick the algorithm.
- These are simple concepts to understand.
- Regression and classification are supervised learning approach that maps an input to an output
- Clustering is a unsupervised learning approach.

# Classification



- Predicts discrete number of values.
- Binary classification — when there is only two classes to predict, usually 1 or 0 values.
- Multi-Class Classification — When there are more than two class labels to predict

# Algorithms for classification

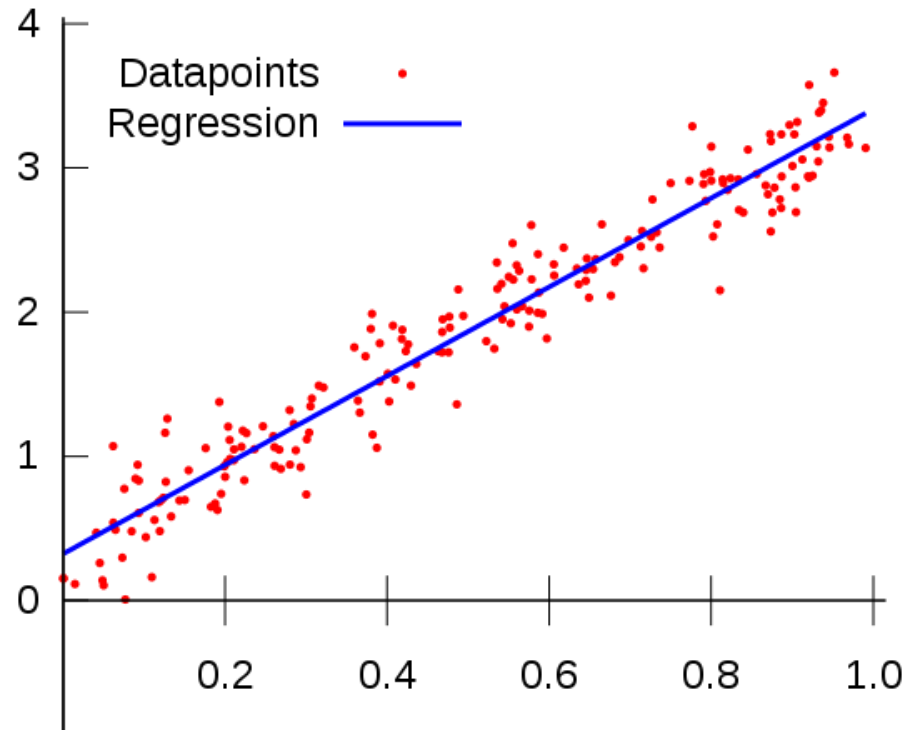
- Decision Trees
- Logistic Regression
- Naive Bayes
- K Nearest Neighbors
- Neural networks

# Application of Classification Algorithms

- Email spam classification
- Bank customers loan pay bank willingness prediction.
- Cancer tumor cells identification.
- Sentiment analysis.
- Drugs classification
- Facial key points detection
- Pedestrians detection in an automotive car driving.

# Regression Problems

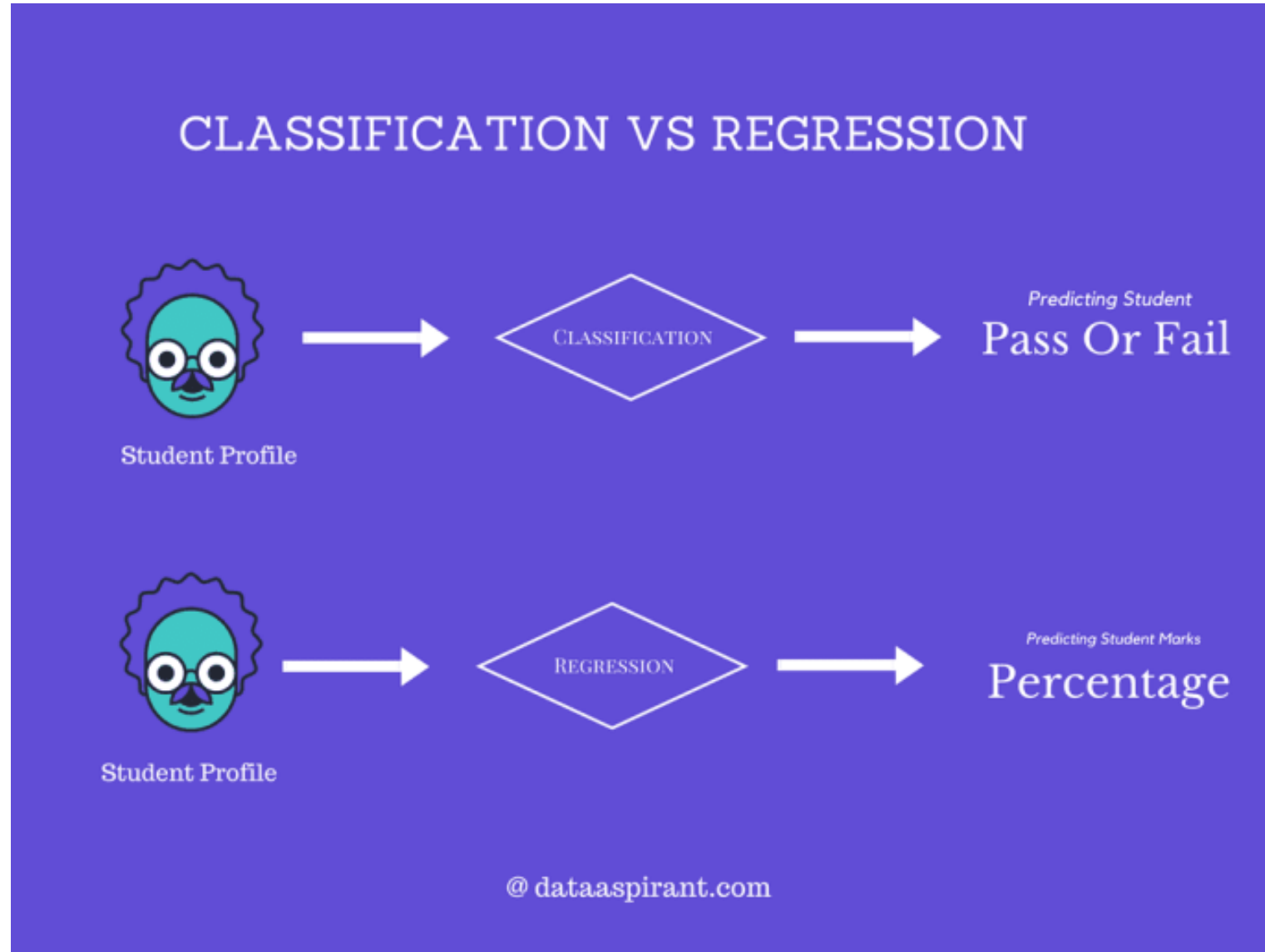
- In regression problems we trying to predict continuous valued output.
- Example: Given a size of the house predict the price(real value).
- Example: Predicting a person's income from their age, education



# Algorithms for regression

- Simple Linear Regression model
- Multivariate Regression algorithm
- Decision Trees
- Random forest

# Classification vs Regression



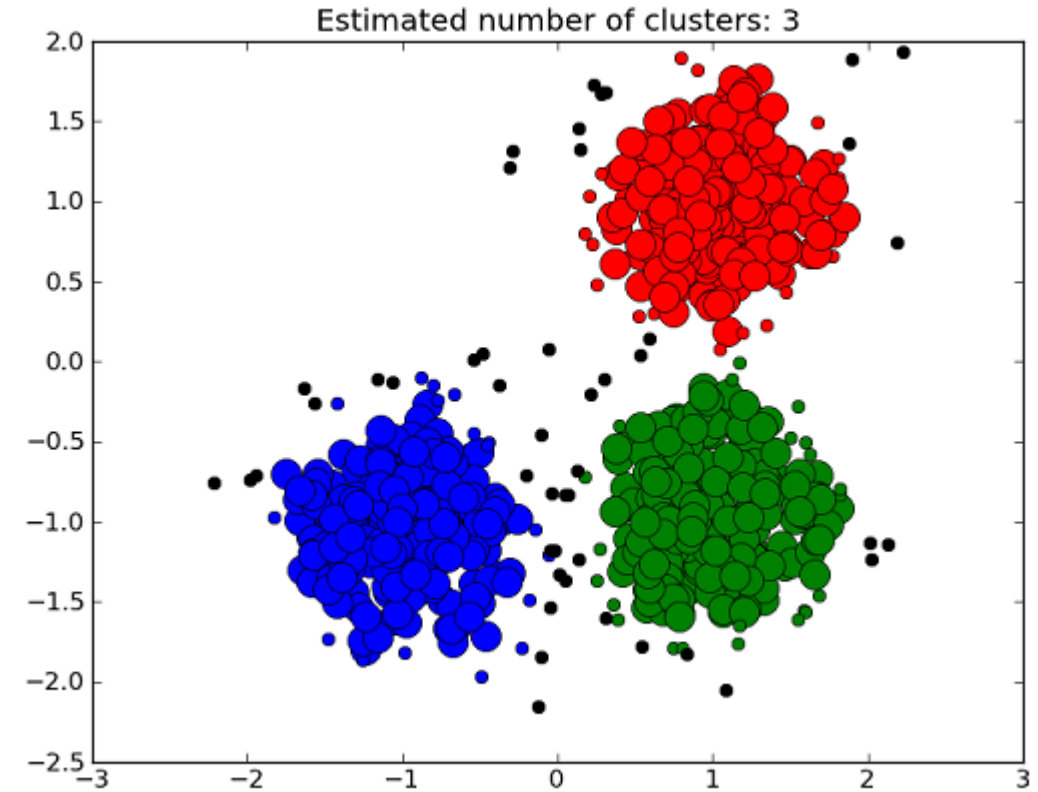
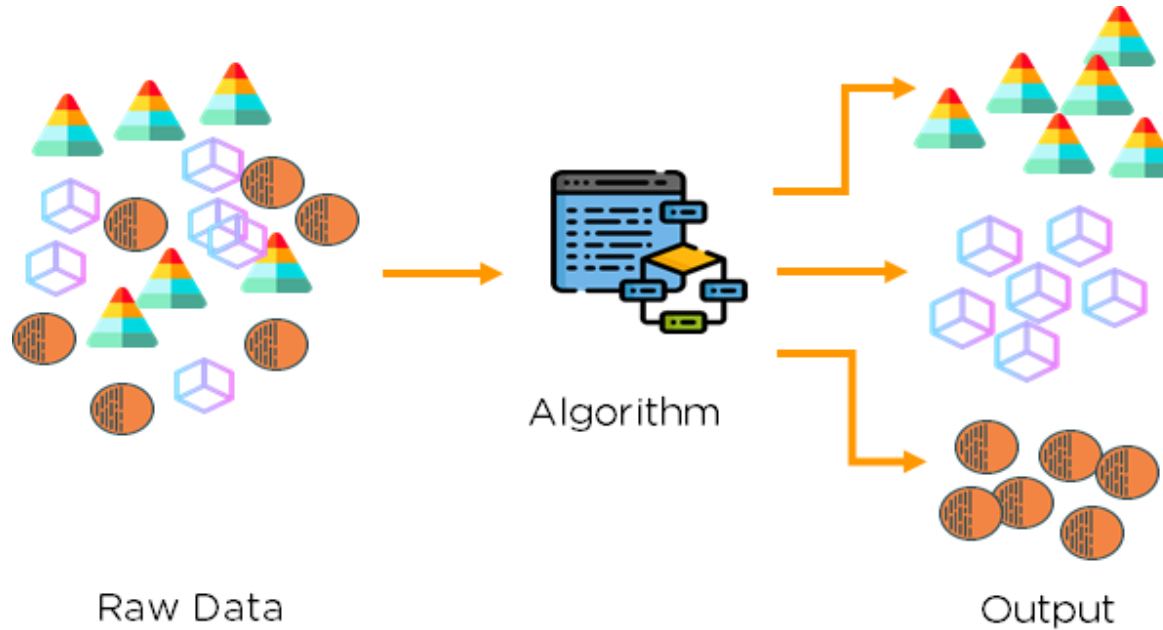


# Clustering

- Clustering is the task of partitioning the dataset into groups, called clusters.
- The goal is to split up the data in such a way that points within each cluster are very similar.
- It determines grouping among unlabeled data.
- Its grouping a set of new data based on similarities amongst them.
- Given a set of data points, we can use a clustering algorithm to classify each data point into a specific group.

# Algorithms for Clustering

- K-Means Clustering
- Hierarchical clustering



# Application of Clustering Algorithms

- Recommender systems
- Anomaly detection
- Grouping of shopping items
- Search result grouping

# Clustering vs Classification

- Classification
  - Used in supervised learning technique where predefined labels are assigned to instances by properties
- Clustering:
  - Used in unsupervised learning where similar instances are grouped, based on their features or properties.

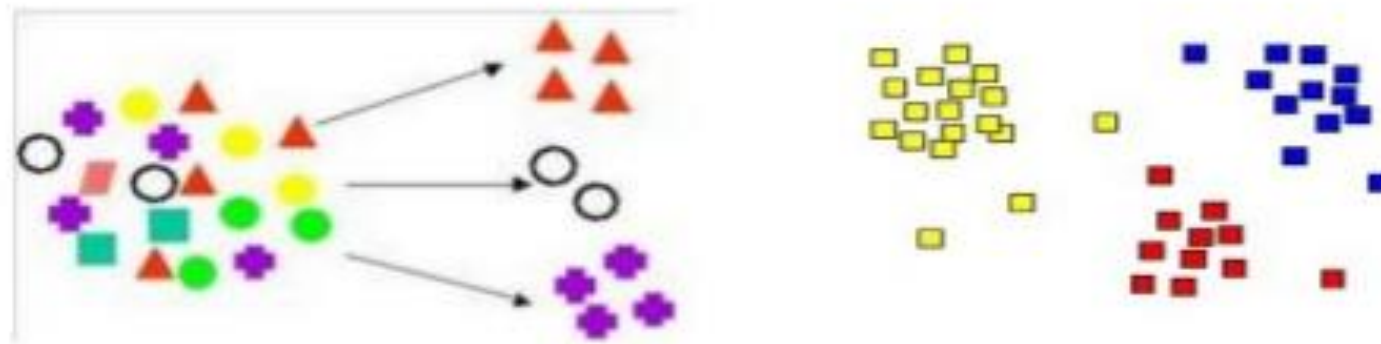
# Clustering vs Classification

<b>BASIS FOR COMPARISON</b>	<b>CLASSIFICATION</b>	<b>CLUSTERING</b>
Involved in	Supervised learning	Unsupervised learning
Training sample	Provided	Not provided
Example	Deciding whether a particular patient record can be associated with a specific disease.	Grouping patient records with similar symptoms without knowing what the symptoms indicate.
Learns	The association between the features of the instance and the class they belong to.	The grouping.

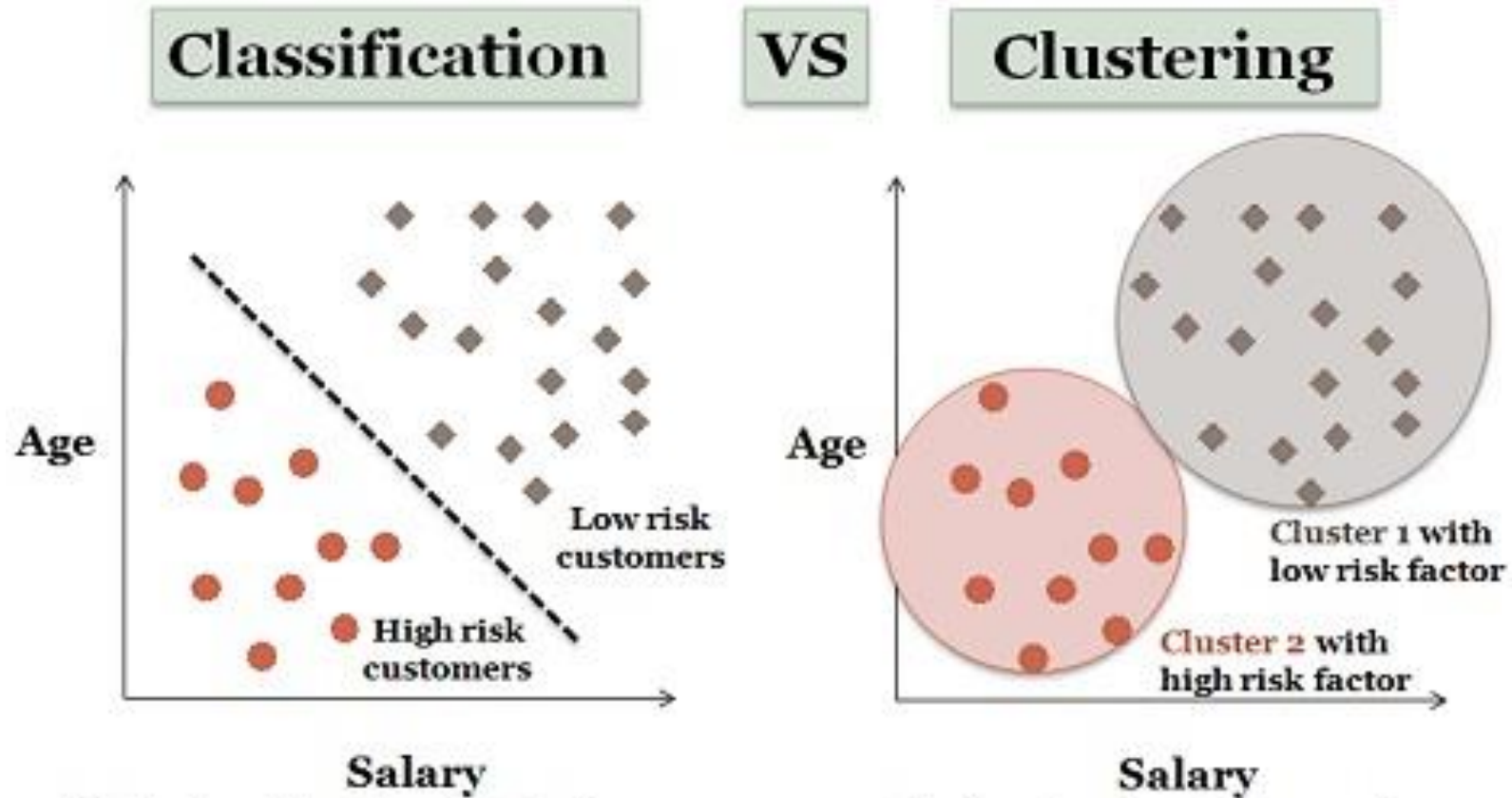
# Clustering vs Classification



## Examples of Clustering



# Clustering vs Classification



Risk classification for the loan payees on the basis of customer salary

Thanks