

# Support Vector Machines

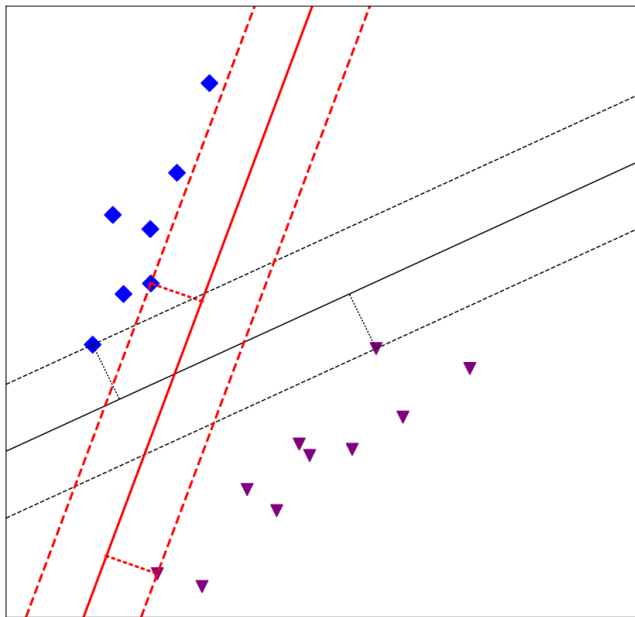
Steve Lopez, Melanie Mejia, Marcello San Agustin

June 2019

# What is SVM?

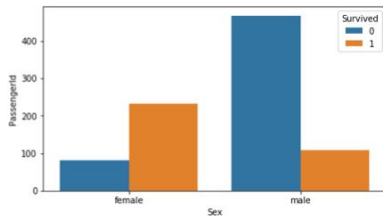
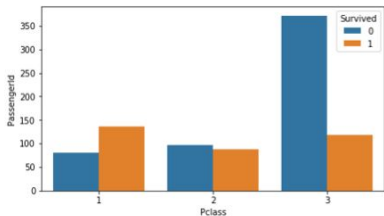
- It is a supervised learning model with learning algorithms that analyze data used for classification and regression analysis.
- A classifier that makes a decision based on finding the optimal support.

# What is SVM?



# Cleaning the Data

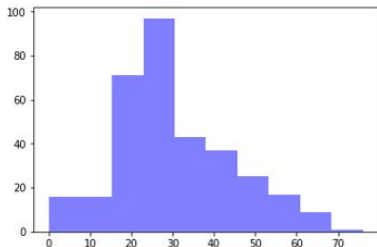
- We began by looking at the data and determining factors that could be important in predicting a passengers fate.



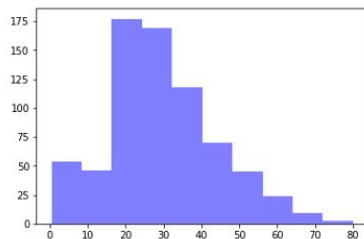
# Cleaning the Data

Then, we determined how to fill in missing data.

Test



Train



# Kaggle Score no.1

- First score: 0.5555502
- Variables used: Age, Sex
- Variables dropped: Name, Cabin, Parch, Pclass, SibSp, Ticket, Fare, PassengerId, Embarked

# How we improved the score

- Considered more variables
- Filled in missing data using the mean of the other data

# Kaggle Score no.2

- Second score: 0.76555
- Variables used: PassengerId, Age, Sex, Pclass
- Variables Dropped: Ticket, Fare, Cabin, Embarked, SibSp, Parch, Name



# Ideas for further improving

- Group data such as price of fare or age
- Group passengers into families or single riders
- Run more kernels and find best parameters such as gamma and C value

# References

- ① Chapter 17. Support Vector Machines. Getting Started in Machine Learning, by Bruce Shapiro, B. E. Shapiro, 2019, pp. 185190.
- ② Gandhi, Rohith. Support Vector Machine Introduction to Machine Learning Algorithms. Towardsdatascience.com, 7 June 2018, [towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47](https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47).
- ③ SVC Beginner. Kaggle, 2017, [www.kaggle.com/pratikjayarao/svc-beginner](https://www.kaggle.com/pratikjayarao/svc-beginner).
- ④ Titanic: Machine Learning from Disaster - SVC. Kaggle, 2018, [www.kaggle.com/jackiy/titanic-machine-learning-from-disaster-svc](https://www.kaggle.com/jackiy/titanic-machine-learning-from-disaster-svc).

# Acknowledgments

Award no.1842386

We would like to thank NSF for the support for the Data Science Program, Career Support and Connections to Industry Program. In addition, we are grateful for the faculty and staff that partake in the program at CSUN.

