

# Çağla Taşdemir



- Istanbul Technical University
  - B.Sc & M.Sc Telecommunications Engineering



- Stony Brook University
  - Ph.D, Electrical Engineering
    - Monte Carlo methods in signal processing
    - Parameter estimation
    - Model selection
    - Bayesian inference





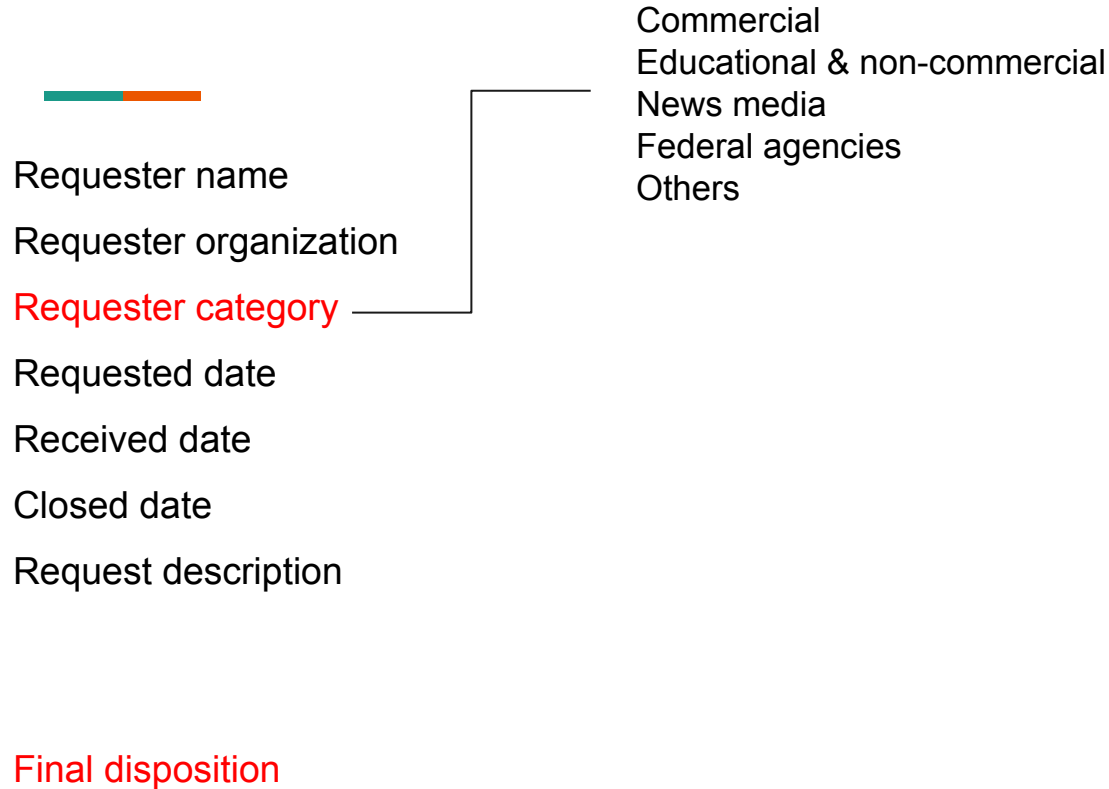
**Grant myFOIA:  
Will your Freedom of Information  
Act Request be approved?**

# About FOIA




- Since 1967, the Freedom of Information Act (FOIA) has provided the public the right to request access to records from any federal agency.
- Predicting **the request success (granted/denied)** of a FOIA request would be useful for both individuals and organizations.
- This prediction can help journalists decide whether to hold a story while waiting to hear back.

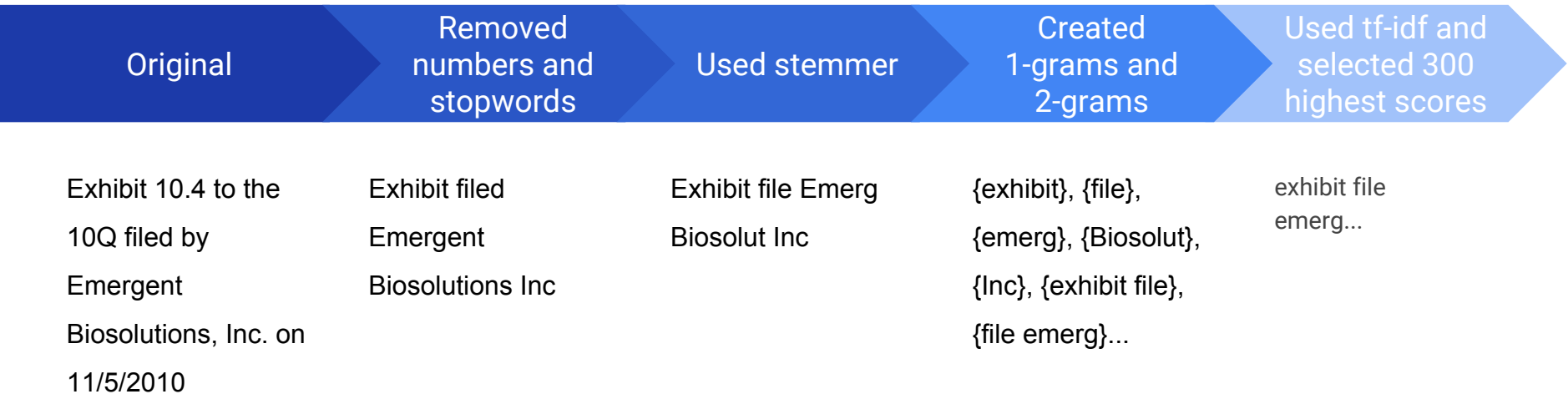
# Dataset - Securities and Exchange Commission



# Dataset - Securities and Exchange Commission

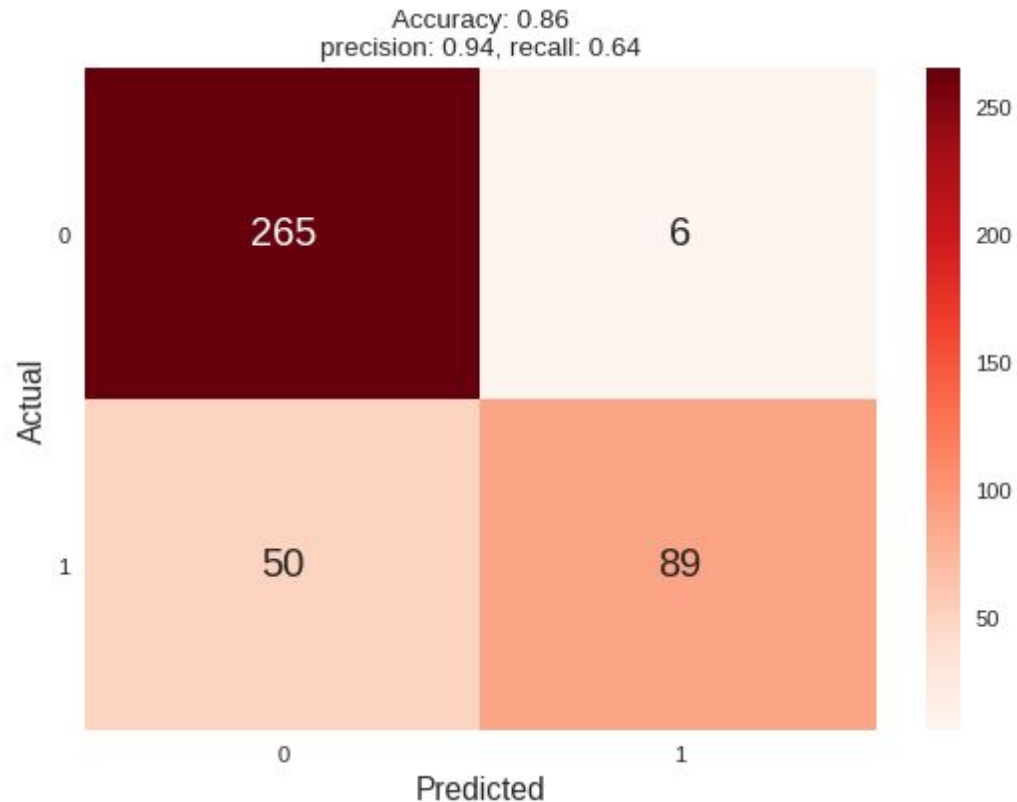
	
Requester name	
Requester organization	
Requester category	News Media → ~15%
Requested date	2015-06-22 2015-07-12
Received date	
Closed date	Exhibit 10.4 to the 10Q filed by Emergent Biosolutions, Inc. on 11/5/2010 The Bank of New England Corporation
Request description	Residential mortgage request John Lynch FOIA Request No. 09-02560-FOIA
Final disposition	Denied Granted

# Request Description - NLP

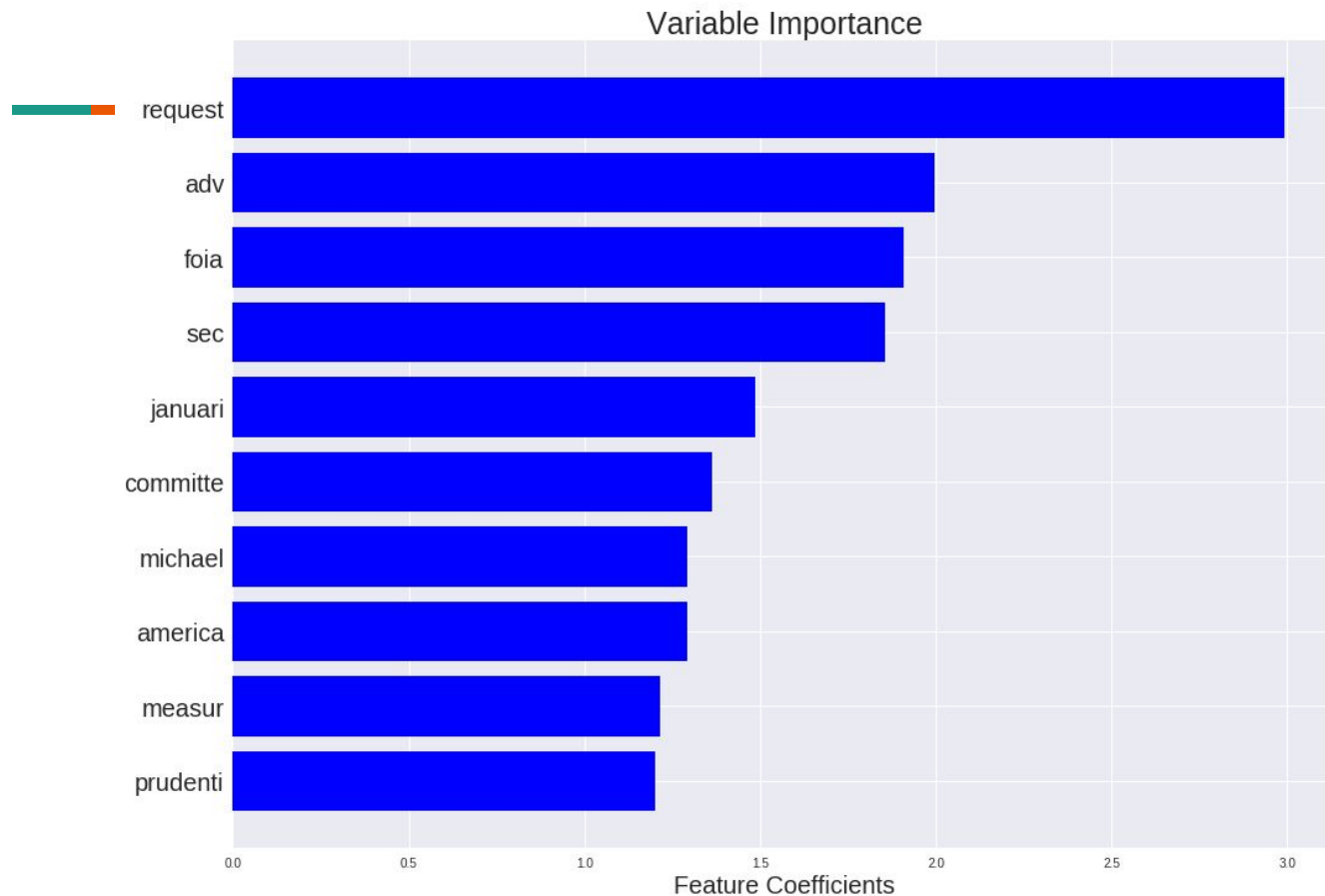


# Classification : Logistic Regression

- Dataset with 1637 requests
  - Training set
    - 418 granted
    - 809 denied
  - Test set
    - 139 granted
    - 271 denied
- Logistic regression with L2 penalty
  - Small dataset, many features
  - Linear classifier
- Imbalanced dataset
  - Updated class weights

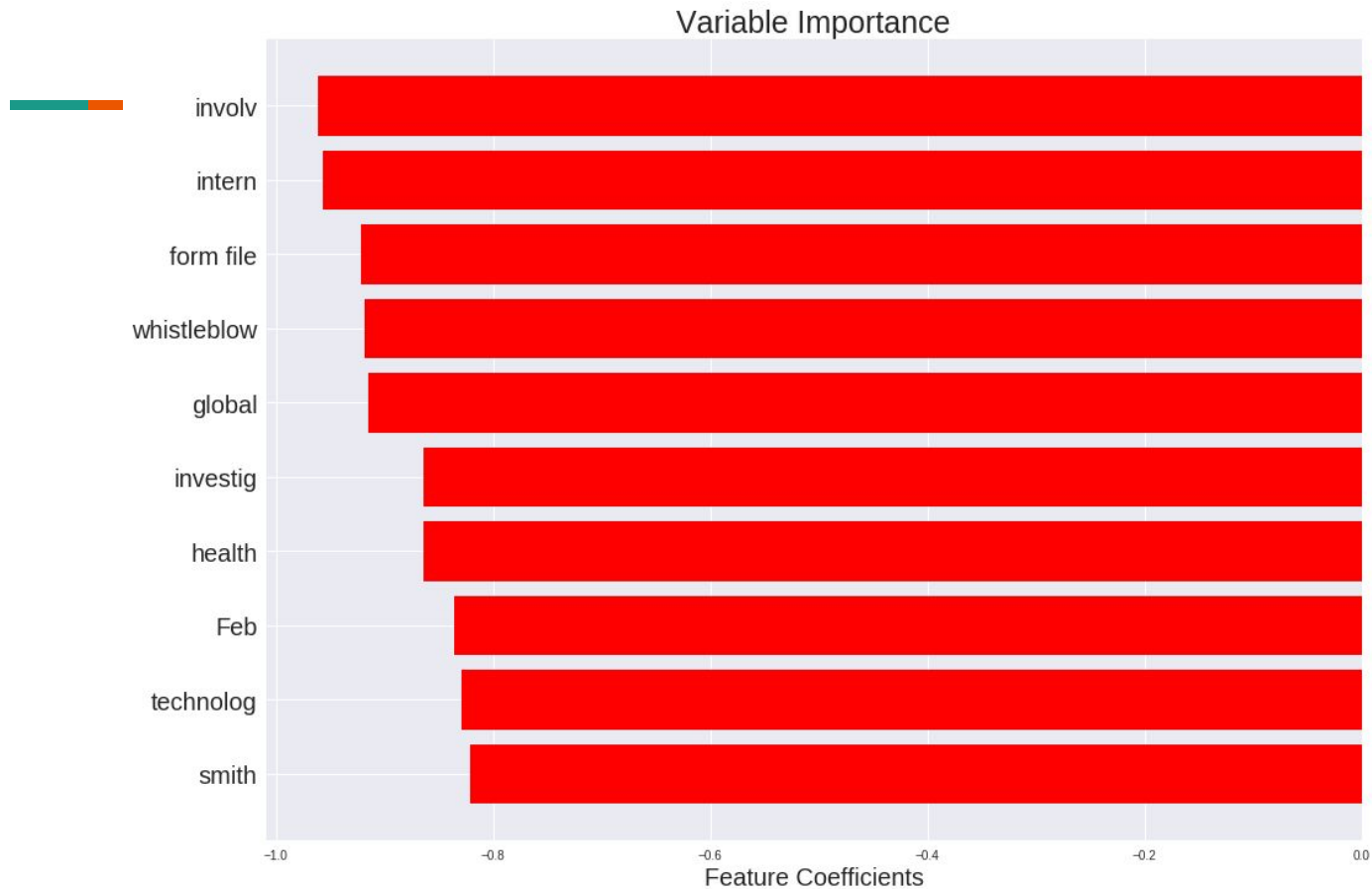


# Feature Importances - positive coefficients





# Feature Importances - negative coefficients



# Prediction results:



Correctly classified requests	Actual	Prediction
applications for <b>whistleblower</b> awards	<b>Denied</b>	<b>Denied</b>
<b>FOIA request</b> 16-02523-FOIA filed by the Democratic National <b>Committee</b>	<b>Granted</b>	<b>Granted</b>

Misclassified requests	Actual	Prediction
communications from the <b>SEC</b> to the California Public Employees' Retirement System	<b>Denied</b>	<b>Granted</b>
Net Element <b>Intern</b> ational	<b>Granted</b>	<b>Denied</b>

# Learning Curve - Logistic Regression





**Thank you!**  
**Grant my FOIA**

# Çağla Taşdemir



- Istanbul Technical University
  - B.Sc & M.Sc Telecommunications Engineering

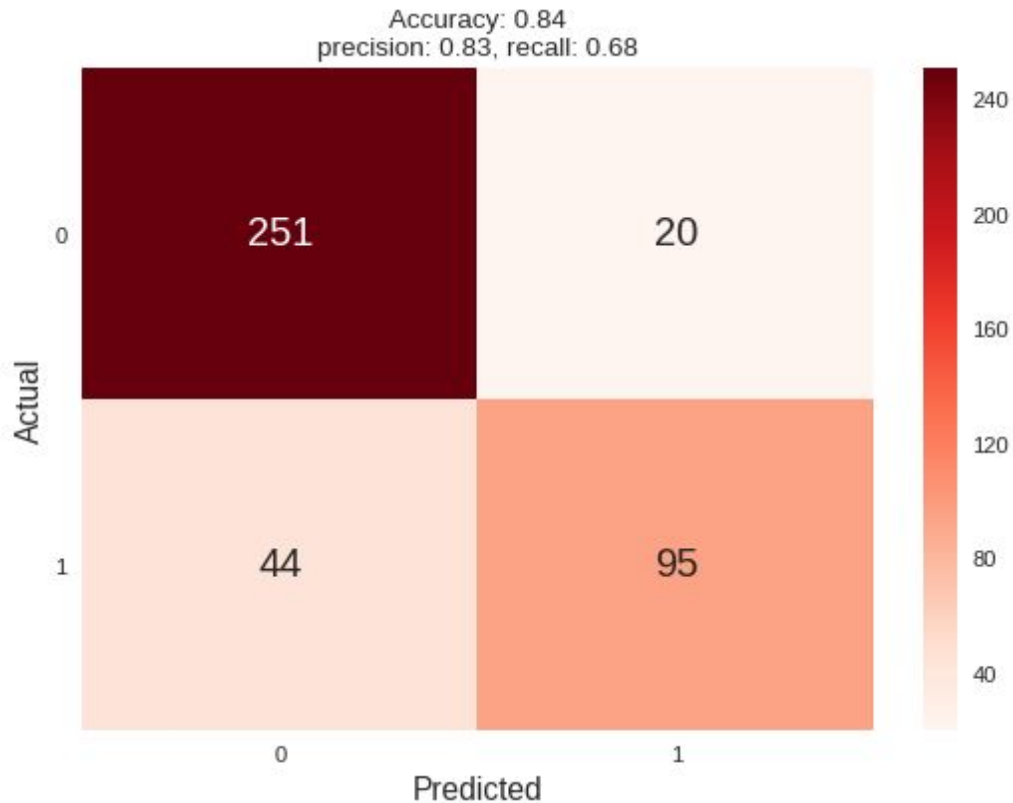


- Stony Brook University
  - Ph.D, Electrical Engineering
    - Monte Carlo methods in signal processing
    - Parameter estimation
    - Model selection
    - Bayesian inference

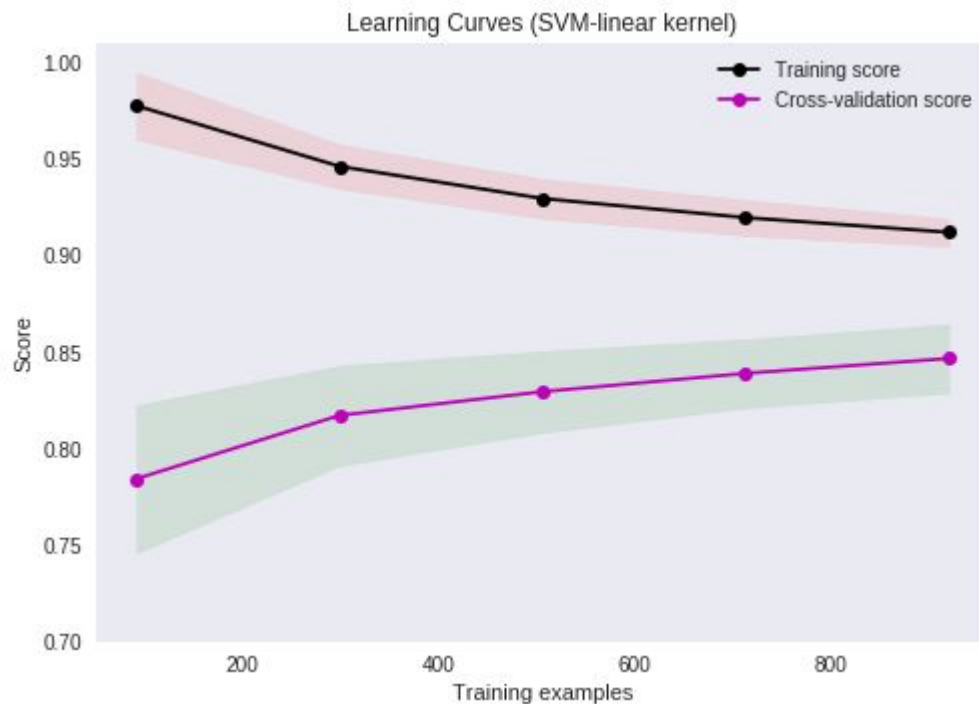


# Classification : SVM with linear kernel

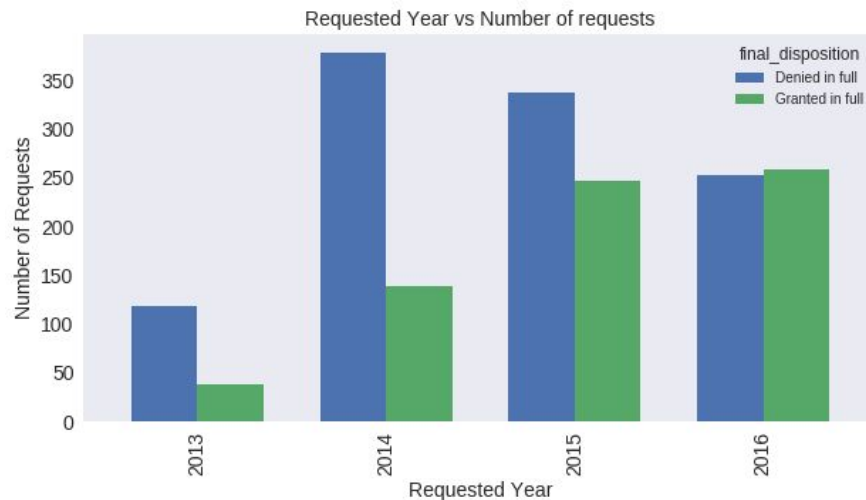
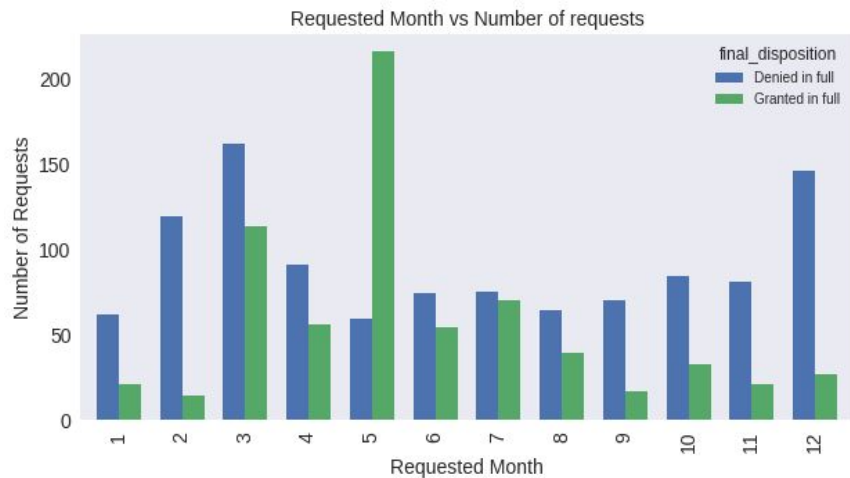
- Dataset with 1637 requests
- Training set with
  - 418 granted and 809 denied requests.
- Test set with
  - 139 granted 271 denied requests.



# Learning curve - SVM with linear kernel



# Requested Month & Year





# Classification : Random forest

- Dataset with 1637 requests
- Training set with
  - 418 granted and 809 denied requests.
- Test set with
  - 139 granted 271 denied requests.

