FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Crime Category Prediction of Cities Using an Ensemble of Various Classifiers

January 11, 2018

Hardy Hambsch Marius Nolden Jessica Ahring Christian Peters

Motivation

Motivation

> Limited resources regarding fighting crimes

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers

Motivation

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers
- Goal: Help those who distribute the resources

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers
- Goal: Help those who distribute the resources
- Predict criminal category of city

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers
- Goal: Help those who distribute the resources
- Predict criminal category of city
  - Either "High", "Medium" or "Low"

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers
- Goal: Help those who distribute the resources
- Predict criminal category of city
  - Either "High", "Medium" or "Low"
  - Depends on the crime rate

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers
- Goal: Help those who distribute the resources
- Predict criminal category of city
  - Either "High", "Medium" or "Low"
  - Depends on the crime rate
- Makes it possible to simulate how certain influences impact the crime category

Motivation

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

- > Limited resources regarding fighting crimes
  - > Limited amounts of money
  - > Limited number of police officers
- > Goal: Help those who distribute the resources
- > Predict criminal category of city
  - > Either "High", "Medium" or "Low"
  - > Depends on the crime rate
- > Makes it possible to simulate how certain influences impact the crime category
  - > Unemployment

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers
- Goal: Help those who distribute the resources
- Predict criminal category of city
  - Either "High", "Medium" or "Low"
  - Depends on the crime rate
- Makes it possible to simulate how certain influences impact the crime category
  - Unemployment
  - Median household income

Motivation

- Limited resources regarding fighting crimes
  - Limited amounts of money
  - Limited number of police officers
- Goal: Help those who distribute the resources
- Predict criminal category of city
  - Either "High", "Medium" or "Low"
  - Depends on the crime rate
- Makes it possible to simulate how certain influences impact the crime category
  - Unemployment
  - Median household income
  - ...

Method

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Method

1. Select most important features of a city

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   > k-Nearest Neighbors

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   > k-Nearest Neighbors
   > Naïve Bayes

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   > k-Nearest Neighbors
   > Naïve Bayes
   > Decision Tree

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   - k-Nearest Neighbors
   - Naïve Bayes
   - Decision Tree
   - Neural Network

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   > k-Nearest Neighbors
   > Naïve Bayes
   > Decision Tree
   > Neural Network
3. Combine the classifiers into an ensemble

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   > k-Nearest Neighbors
   > Naïve Bayes
   > Decision Tree
   > Neural Network
3. Combine the classifiers into an ensemble
   > Majority vote

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   - k-Nearest Neighbors
   - Naïve Bayes
   - Decision Tree
   - Neural Network
3. Combine the classifiers into an ensemble
   - Majority vote
4. Evaluate the results

Method

1. Select most important features of a city
2. Train basic classifiers using WEKA
   > k-Nearest Neighbors
   > Naïve Bayes
   > Decision Tree
   > Neural Network
3. Combine the classifiers into an ensemble
   > Majority vote
4. Evaluate the results
   > Accuracy, Precision, Recall, F-Measure

Dataset

Dataset

> Union of three subsets containing real crime data

Dataset

> Union of three subsets containing real crime data
>   > 1990 US Census

Dataset

> Union of three subsets containing real crime data
  > 1990 US Census
  > 1990 US LEMAS survey

Dataset

- Union of three subsets containing real crime data
  - 1990 US Census
  - 1990 US LEMAS survey
  - 1995 FBI UCR

Dataset

- Union of three subsets containing real crime data
  - 1990 US Census
  - 1990 US LEMAS survey
  - 1995 FBI UCR
- Contains 1994 instances (cities) and 128 attributes

Dataset

- Union of three subsets containing real crime data
  - 1990 US Census
  - 1990 US LEMAS survey
  - 1995 FBI UCR
- Contains 1994 instances (cities) and 128 attributes
- 13 features selected for classification

Dataset

- Union of three subsets containing real crime data
  - 1990 US Census
  - 1990 US LEMAS survey
  - 1995 FBI UCR
- Contains 1994 instances (cities) and 128 attributes
- 13 features selected for classification
  - Each feature is continuous, normalized and has no missing values

Dataset

- Union of three subsets containing real crime data
  - 1990 US Census
  - 1990 US LEMAS survey
  - 1995 FBI UCR
- Contains 1994 instances (cities) and 128 attributes
- 13 features selected for classification
  - Each feature is continuous, normalized and has no missing values
  - Details of feature selection approach discussed later ...

Dataset

- Union of three subsets containing real crime data
  - 1990 US Census
  - 1990 US LEMAS survey
  - 1995 FBI UCR
- Contains 1994 instances (cities) and 128 attributes
- 13 features selected for classification
  - Each feature is continuous, normalized and has no missing values
  - Details of feature selection approach discussed later ...
- The criminal category ("High", "Medium", "Low") was added based on the crime rate

Dataset

- Union of three subsets containing real crime data
  - 1990 US Census
  - 1990 US LEMAS survey
  - 1995 FBI UCR
- Contains 1994 instances (cities) and 128 attributes
- 13 features selected for classification
  - Each feature is continuous, normalized and has no missing values
  - Details of feature selection approach discussed later ...
- The criminal category ("High", "Medium", "Low") was added based on the crime rate
  - Details discussed later ...

List of selected features

List of selected features

> percentage african american

List of selected features

- percentage african american
- percentage caucasian

List of selected features

- percentage african american
- percentage caucasian
- median household income

List of selected features

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income

List of selected features

- > percentage african american
- > percentage caucasian
- > median household income
- > percentage of households with investment / rent income
- > percentage of households with public assistance income

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

List of selected features

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income
- percentage of households with public assistance income
- percentage below the poverty level

List of selected features

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income
- percentage of households with public assistance income
- percentage below the poverty level
- percentage of people 25 and over with less than 9th grade education

List of selected features

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income
- percentage of households with public assistance income
- percentage below the poverty level
- percentage of people 25 and over with less than 9th grade education
- unemployment rate

List of selected features

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income
- percentage of households with public assistance income
- percentage below the poverty level
- percentage of people 25 and over with less than 9th grade education
- unemployment rate
- divorce rate

List of selected features

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income
- percentage of households with public assistance income
- percentage below the poverty level
- percentage of people 25 and over with less than 9th grade education
- unemployment rate
- divorce rate
- percentage of kids living with two parents

List of selected features

> percentage african american
> percentage caucasian
> median household income
> percentage of households with investment / rent income
> percentage of households with public assistance income
> percentage below the poverty level
> percentage of people 25 and over with less than 9th grade education
> unemployment rate
> divorce rate
> percentage of kids living with two parents
> percentage of kids born to never married

List of selected features

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income
- percentage of households with public assistance income
- percentage below the poverty level
- percentage of people 25 and over with less than 9th grade education
- unemployment rate
- divorce rate
- percentage of kids living with two parents
- percentage of kids born to never married
- percentage of people who don't speak English well

List of selected features

- percentage african american
- percentage caucasian
- median household income
- percentage of households with investment / rent income
- percentage of households with public assistance income
- percentage below the poverty level
- percentage of people 25 and over with less than 9th grade education
- unemployment rate
- divorce rate
- percentage of kids living with two parents
- percentage of kids born to never married
- percentage of people who don't speak English well
- percentage of people in owner occupied households

List of selected features

- > percentage african american
- > percentage caucasian
- > median household income
- > percentage of households with investment / rent income
- > percentage of households with public assistance income
- > percentage below the poverty level
- > percentage of people 25 and over with less than 9th grade education
- > unemployment rate
- > divorce rate
- > percentage of kids living with two parents
- > percentage of kids born to never married
- > percentage of people who don't speak English well
- > percentage of people in owner occupied households
- > **crime category** (added by us)

Literature Review

Literature Review

> The authors of [Kha13] developed a similar approach

Literature Review

- The authors of [Kha13] developed a similar approach
  - Same dataset

Literature Review

- The authors of [Kha13] developed a similar approach
  - Same dataset
  - Only Naïve Bayes and Decision Tree

Literature Review

> The authors of [Kha13] developed a similar approach
  > Same dataset
  > Only Naïve Bayes and Decision Tree
  > Feature selection based on subjective choice

Literature Review

> The authors of [Kha13] developed a similar approach
  > Same dataset
  > Only Naïve Bayes and Decision Tree
  > Feature selection based on subjective choice
  > No details disclosed regarding their method of assigning the class labels

Literature Review

> The authors of [Kha13] developed a similar approach
  > Same dataset
  > Only Naïve Bayes and Decision Tree
  > Feature selection based on subjective choice
  > No details disclosed regarding their method of assigning the class labels

> A different approach is illustrated in [Jon03]

Literature Review

- The authors of [Kha13] developed a similar approach
  - Same dataset
  - Only Naïve Bayes and Decision Tree
  - Feature selection based on subjective choice
  - No details disclosed regarding their method of assigning the class labels
- A different approach is illustrated in [Jon03]
  - Goal: predict crime hotspots in the US

Literature Review

- The authors of [Kha13] developed a similar approach
  - Same dataset
  - Only Naïve Bayes and Decision Tree
  - Feature selection based on subjective choice
  - No details disclosed regarding their method of assigning the class labels
- A different approach is illustrated in [Jon03]
  - Goal: predict crime hotspots in the US
  - They first mapped unique crime records to the corresponding locations of the US using geographical coordinates

Literature Review

- The authors of [Kha13] developed a similar approach
  - Same dataset
  - Only Naïve Bayes and Decision Tree
  - Feature selection based on subjective choice
  - No details disclosed regarding their method of assigning the class labels
- A different approach is illustrated in [Jon03]
  - Goal: predict crime hotspots in the US
  - They first mapped unique crime records to the corresponding locations of the US using geographical coordinates
  - Their next step was to find clusters

Literature Review

- The authors of [Kha13] developed a similar approach
  - Same dataset
  - Only Naïve Bayes and Decision Tree
  - Feature selection based on subjective choice
  - No details disclosed regarding their method of assigning the class labels
- A different approach is illustrated in [Jon03]
  - Goal: predict crime hotspots in the US
  - They first mapped unique crime records to the corresponding locations of the US using geographical coordinates
  - Their next step was to find clusters
  - The prediction of crime activity within the clusters was performed by an artificial neural network

Data Preparation

Data Preparation

> Before the classification can be done, a few steps are required on the raw data

Data Preparation

> Before the classification can be done, a few steps are required on the raw data

> Original dataset does not contain the class labels ("High", "Medium", "Low")

Data Preparation

- Before the classification can be done, a few steps are required on the raw data
- Original dataset does not contain the class labels ("High", "Medium", "Low")
  - How to assign each instance a crime category based on its crime rate?

Data Preparation

- Before the classification can be done, a few steps are required on the raw data
- Original dataset does not contain the class labels ("High", "Medium", "Low")
  - How to assign each instance a crime category based on its crime rate?
  - How to find the percentage boundaries between the classes?

Data Preparation

- Before the classification can be done, a few steps are required on the raw data
- Original dataset does not contain the class labels ("High", "Medium", "Low")
  - How to assign each instance a crime category based on its crime rate?
  - How to find the percentage boundaries between the classes?
- Curse of dimensionality

Data Preparation

- Before the classification can be done, a few steps are required on the raw data
- Original dataset does not contain the class labels ("High", "Medium", "Low")
  - How to assign each instance a crime category based on its crime rate?
  - How to find the percentage boundaries between the classes?
- Curse of dimensionality
  - Huge number of attributes (128 per city)

Data Preparation

- Before the classification can be done, a few steps are required on the raw data
- Original dataset does not contain the class labels ("High", "Medium", "Low")
  - How to assign each instance a crime category based on its crime rate?
  - How to find the percentage boundaries between the classes?
- Curse of dimensionality
  - Huge number of attributes (128 per city)
  - Select only the most significant

Data Preparation

- Before the classification can be done, a few steps are required on the raw data
- Original dataset does not contain the class labels ("High", "Medium", "Low")
  - How to assign each instance a crime category based on its crime rate?
  - How to find the percentage boundaries between the classes?
- Curse of dimensionality
  - Huge number of attributes (128 per city)
  - Select only the most significant
  - Reduce the dimensionality of the feature space

Assignment of Class Labels

Assignment of Class Labels

> Divide the attribute "crime rate" into different groups

Assignment of Class Labels

- Divide the attribute "crime rate" into different groups
  - Initial question: How many groups?

Assignment of Class Labels

> Divide the attribute "crime rate" into different groups
  > Initial question: How many groups?
  > What are the boundaries between the groups?

Assignment of Class Labels

- Divide the attribute "crime rate" into different groups
  - Initial question: How many groups?
  - What are the boundaries between the groups?
- Use clustering to solve this problem

Assignment of Class Labels

- Divide the attribute "crime rate" into different groups
  - Initial question: How many groups?
  - What are the boundaries between the groups?
- Use clustering to solve this problem
  - Simple k-Means clustering algorithm

Assignment of Class Labels

- Divide the attribute "crime rate" into different groups
  - Initial question: How many groups?
  - What are the boundaries between the groups?
- Use clustering to solve this problem
  - Simple k-Means clustering algorithm
  - Only cluster the crime rate (i.e. only one dimension)

Assignment of Class Labels

- Divide the attribute "crime rate" into different groups
  - Initial question: How many groups?
  - What are the boundaries between the groups?
- Use clustering to solve this problem
  - Simple k-Means clustering algorithm
  - Only cluster the crime rate (i.e. only one dimension)
  - Test different k values (different amounts of groups)

Assignment of Class Labels

- Divide the attribute "crime rate" into different groups
  - Initial question: How many groups?
  - What are the boundaries between the groups?
- Use clustering to solve this problem
  - Simple k-Means clustering algorithm
  - Only cluster the crime rate (i.e. only one dimension)
  - Test different k values (different amounts of groups)
  - Find the boundaries between the groups

Assignment of Class Labels

FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

- Divide the attribute "crime rate" into different groups
  - Initial question: How many groups?
  - What are the boundaries between the groups?
- Use clustering to solve this problem
  - Simple k-Means clustering algorithm
  - Only cluster the crime rate (i.e. only one dimension)
  - Test different k values (different amounts of groups)
  - Find the boundaries between the groups
- Each cluster will be assigned a label (e.g. "High", "Medium", "Low",. . . crime rate)

Assignment of Class Labels

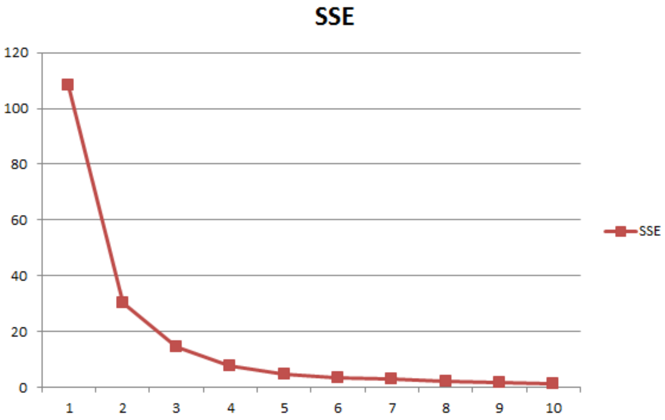

Figure: Sum of the squared errors for different choices of k

Assignment of Class Labels

Assignment of Class Labels

> By using the elbow-method, $k = 3$ is the best choice

Assignment of Class Labels

- By using the elbow-method, $k = 3$ is the best choice
  - This means to divide the crime rate into three groups ("High", "Medium", "Low")

Assignment of Class Labels

- By using the elbow-method, $k = 3$ is the best choice
  - This means to divide the crime rate into three groups ("High", "Medium", "Low")
- The percentage boundaries are extracted from the clusters

Assignment of Class Labels

> By using the elbow-method, $k = 3$ is the best choice
  > This means to divide the crime rate into three groups ("High", "Medium", "Low")

> The percentage boundaries are extracted from the clusters
  > Low: $[0\%; 22\%]$

Assignment of Class Labels

- By using the elbow-method, $k = 3$ is the best choice
  - This means to divide the crime rate into three groups ("High", "Medium", "Low")
- The percentage boundaries are extracted from the clusters
  - Low: $[0\%; 22\%]$
  - Medium: $(22\%; 56\%]$

Assignment of Class Labels

- By using the elbow-method, $k = 3$ is the best choice
  - This means to divide the crime rate into three groups ("High", "Medium", "Low")
- The percentage boundaries are extracted from the clusters
  - Low: $[0\%; 22\%]$
  - Medium: $(22\%; 56\%]$
  - High: $(56\%; 100\%]$

Assignment of Class Labels

> By using the elbow-method, $k = 3$ is the best choice
  > This means to divide the crime rate into three groups ("High", "Medium", "Low")

> The percentage boundaries are extracted from the clusters
  > Low: $[0\%; 22\%]$
  > Medium: $(22\%; 56\%]$
  > High: $(56\%; 100\%]$

> Depending on the group it falls into, each instance (city) can now be assigned a label

Feature Selection

Feature Selection

> Two-step approach:

Feature Selection

> Two-step approach:
1. Create a ranking by judgement of significance

Feature Selection

- Two-step approach:
  1. Create a ranking by judgement of significance
     - Discuss each attribute

Feature Selection

> Two-step approach:
  1. Create a ranking by judgement of significance
     > Discuss each attribute
     > Agree if it is included or not

Feature Selection

> Two-step approach:
1. Create a ranking by judgement of significance
   > Discuss each attribute
   > Agree if it is included or not
2. Generate a ranking based on the correlation between each attribute and the class labels

Feature Selection

> Two-step approach:
1. Create a ranking by judgement of significance
   > Discuss each attribute
   > Agree if it is included or not
2. Generate a ranking based on the correlation between each attribute and the class labels

> Compare the judgement ranking with the correlation ranking

Feature Selection

- Two-step approach:
  1. Create a ranking by judgement of significance
     - Discuss each attribute
     - Agree if it is included or not
  2. Generate a ranking based on the correlation between each attribute and the class labels
- Compare the judgement ranking with the correlation ranking
- Add some features with high correlation and remove some with low correlation

Feature Selection

> Two-step approach:
1. Create a ranking by judgement of significance
    > Discuss each attribute
    > Agree if it is included or not
2. Generate a ranking based on the correlation between each attribute and the class labels

> Compare the judgement ranking with the correlation ranking
> Add some features with high correlation and remove some with low correlation
    > E.g. the attribute "population" was removed because of a correlation near zero

Experimental Results

Approaches and Settings

Experimental Results

Approaches and Settings

> k-Nearest Neighbors

Experimental Results

Approaches and Settings

> k-Nearest Neighbors
> Naïve Bayes

Experimental Results

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

Approaches and Settings

> k-Nearest Neighbors
> Naïve Bayes
> Decision Tree

Experimental Results

Approaches and Settings

> k-Nearest Neighbors
> Naïve Bayes
> Decision Tree
> Neural Network

Experimental Results

Approaches and Settings

- k-Nearest Neighbors
- Naïve Bayes
- Decision Tree
- Neural Network
- Ensemble

k-Nearest Neighbors

k-Nearest Neighbors

> parameter search on a leave-one-out cross validation

k-Nearest Neighbors

> parameter search on a leave-one-out cross validation
> k=30 best (accuracy)

k-Nearest Neighbors

> parameter search on a leave-one-out cross validation
> k=30 best (accuracy)
> distance measure = euclidean distance

k-Nearest Neighbors

> parameter search on a leave-one-out cross validation
> k=30 best (accuracy)
> distance measure = euclidean distance
> no distance weighting => better results

Naïve Bayes

> Useful in may areas (text classification, medical diagnosis, . . .)

Naïve Bayes

- Useful in may areas (text classification, medical diagnosis, ...)
- Still employed although the features are obviously not independent

Naïve Bayes

> Useful in may areas (text classification, medical diagnosis, ...)
> Still employed although the features are obviously not independent
>   > For example "median income" is clearly related to "number of people below the poverty level"

Decision Tree

Decision Tree

> pruned J48-decision tree

Decision Tree

- pruned J48-decision tree
- confidence factor $c = 0.045$

Decision Tree

- pruned J48-decision tree
- confidence factor $c = 0.045$
- found by utilizing WEKAs capabilities of linear parameter search

Neural Network

Neural Network

> multilayer perceptron

Neural Network

- multilayer perceptron
- 13 input neurons (features)

Neural Network

- multilayer perceptron
- 13 input neurons (features)
- hidden layer of 17 neurons

Neural Network

- multilayer perceptron
- 13 input neurons (features)
- hidden layer of 17 neurons
- output layer with three neurons (classes)

Neural Network

- multilayer perceptron
- 13 input neurons (features)
- hidden layer of 17 neurons
- output layer with three neurons (classes)
- learning rate: $\alpha = 0.1$

Neural Network

- multilayer perceptron
- 13 input neurons (features)
- hidden layer of 17 neurons
- output layer with three neurons (classes)
- learning rate: $\alpha = 0.1$
- momentum: $m = 0.1$ (backpropagation algorithm)

Neural Network

- multilayer perceptron
- 13 input neurons (features)
- hidden layer of 17 neurons
- output layer with three neurons (classes)
- learning rate: $\alpha = 0.1$
- momentum: $m = 0.1$ (backpropagation algorithm)
- linear search for parameters: hidden neurons, learning rate and momentum

Ensemble

Ensemble

> Merge all models into an ensemble

Ensemble

> Merge all models into an ensemble

> Goal: Combine the strengths of the different classifiers

Ensemble

- Merge all models into an ensemble
- Goal: Combine the strengths of the different classifiers
- Choice of ensemble was found by conducting a majority voting

Results

> The results of each classifier were evaluated using 10-fold cross validation

Results

- The results of each classifier were evaluated using 10-fold cross validation
- Common evaluation metrics allow for comparison
  - Accuracy, Precision, Recall, F-Measure

Results

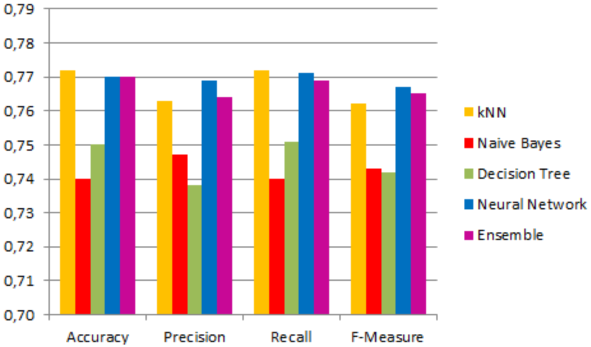

Figure: Evaluation metrics of the different classifiers

Results

Table: Confusion matrix of the ensemble

classified →	**Low**	**Medium**	**High**	Total
Low	1125	126	8	1259
Medium	160	305	57	522
High	15	95	103	213
Total	1300	526	168	1994

Recall

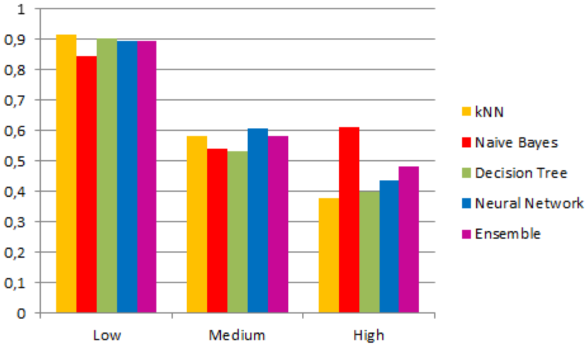

Figure: Recall of the different classifiers

Discussion and Outlook

Discussion and Outlook

> Each classifier has its strengths and weaknesses

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"

Discussion and Outlook

> Each classifier has its strengths and weaknesses
  > k-Nearest Neighbors excels in accuracy and overall recall
  > Neural Network scores best in terms of precision and f-measure
  > Naïve Bayes has extraordinary recall with respect to class "High"
> Contrary to what we initially thought – the ensemble was merely average

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category
- How can that be?

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category
- How can that be?
  - Not only strengths are merged but weaknesses as well

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category
- How can that be?
  - Not only strengths are merged but weaknesses as well
  - Strengths and weaknesses average out

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category
- How can that be?
  - Not only strengths are merged but weaknesses as well
  - Strengths and weaknesses average out
- Can this be improved?

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category
- How can that be?
  - Not only strengths are merged but weaknesses as well
  - Strengths and weaknesses average out
- Can this be improved?
  - In our case, each base learner had an equal weight

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category
- How can that be?
  - Not only strengths are merged but weaknesses as well
  - Strengths and weaknesses average out
- Can this be improved?
  - In our case, each base learner had an equal weight
  - Improve the results by changing the weights

Discussion and Outlook

- Each classifier has its strengths and weaknesses
  - k-Nearest Neighbors excels in accuracy and overall recall
  - Neural Network scores best in terms of precision and f-measure
  - Naïve Bayes has extraordinary recall with respect to class "High"
- Contrary to what we initially thought – the ensemble was merely average
  - Always among the best
  - Never reached top spot in any category
- How can that be?
  - Not only strengths are merged but weaknesses as well
  - Strengths and weaknesses average out
- Can this be improved?
  - In our case, each base learner had an equal weight
  - Improve the results by changing the weights
  - Emphasize strengths of one classifier (e.g. high recall of Naïve Bayes), mitigate weaknesses of another

Discussion and Outlook

Thank you very much for listening!
Any questions left?

J. Andrew Ware Jonathan J. Corcoran Ian D. Wilson. "Predicting the geo-temporal variations of crime and disorder". English. In: *International Journal of Forecasting* 19(4) (Oct. 2003), pp. 623–634. URL: http://www.sciencedirect.com/science/article/pii/S0169207003000955.

Nasim Khanahmadliravi. "An Experimental Study of Classification Algorithms for Crime Prediction". English. In: *Indian Journal of Science and Technology* 6(3) (Mar. 2013), pp. 4219–4225. ISSN: 0974-6846. URL: http://52.172.159.94/index.php/indjst/article/download/31230/27028.