Using Bayesian Networks to Estimating Rainfall Distribution Given Polarimetric Radar Data

Jose Picado
Vahid Ghadakchi
Zahra Iman

Probabilistic Graphical Models
Winter 2015
Introduction

• Rainfall amount prediction problem
  – Using gauges
  – Remote sensing

• Using Bayes Nets
  – Designing the structure
  – Learning the structure

Image courtesy of NOAA
### Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>RR1</th>
<th>RR2</th>
<th>RR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>TimeToEnd</td>
<td>RhoHV</td>
<td>Velocity</td>
<td>Zdr</td>
</tr>
<tr>
<td>DistanceToRadar</td>
<td>RR2</td>
<td>MassWeightedMean</td>
<td>LogWaterVolume</td>
</tr>
<tr>
<td>Composite</td>
<td>RR3</td>
<td>Reflectivity</td>
<td>RadarQualityIndex</td>
</tr>
<tr>
<td>HybridScan</td>
<td>Reflectivity</td>
<td>MassWeightedMean</td>
<td>MassWeightedMean</td>
</tr>
<tr>
<td>HydropmeteorType</td>
<td>ReflectivityQC</td>
<td>MassWeightedSD</td>
<td></td>
</tr>
<tr>
<td>Kdp</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Features with composite values aggregated into a single value
- Introduced feature: Measurements
- Label: Expected – actual amount of rain reported by rain gauge for that hour
Dependencies

• RR1: Rain rate from the HCI-based algorithm
  – HydrometeorType, Zdr, RhoHV, Kdp
• RR2: Rain rate from the Zdr-based algorithm
  – Zdr
• HydrometeorType: rain, snow, big drops, birds,…
  – Based on HCA algorithm with inputs of Zdr, Kdp, RhoHV, Reflectivity
• RadarQualityIndex: value from 0 to 1 (good data)
  – Based on an algorithm proposed in a paper with inputs of Reflectivity, Zdr, Kdp, RhoHV
Approaches

• Bayes network with manually designed structure
• Bayes network with learned structure
Manually designed structure
Learned structure

- Structure learning algorithms:
  - Structure scoring: Bayesian scoring with Dirichlet priors
  - Greedy: Hill Climbing, K2
  - Non-greedy: Simulated Annealing
Learned structure
Results

• Data and Setting
  – Subset of training set
  – Implementation using Weka

• Results
  – Continuous Ranked Probability

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>CRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>0.01096134</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>0.01452578</td>
</tr>
<tr>
<td>Bayes Net - Manually Designed Structure</td>
<td>0.00965175</td>
</tr>
<tr>
<td>Bayes Net - Simulated Annealing</td>
<td>0.00864957</td>
</tr>
</tbody>
</table>
Conclusions

• Bayes networks proved to be an appropriate model:
  – Exploit (in)dependencies between features
• Domain knowledge can be used to build the network structure
• Structure learning is better when data is available and time is not an issue
• Future work:
  – Structure provided by domain expert
  – Aggregations functions for composite-valued features
  – Undirected models
  – Build model with full dataset