Parallel & Scalable Machine Learning

Introduction to Machine Learning Algorithms

Dr. – Ing. Morris Riedel
Adjunct Associated Professor, University of Iceland
Research Group Leader, Juelich Supercomputing Centre, Germany

Dr. Gabriele Cavallaro
Post Doctoral Researcher, Juelich Supercomputing Centre, Germany

LECTURE 6

Classification Applications, Challenges, and Solutions

January 16th, 2018
JSC, Germany
Review of Lecture 5

- **Supervised Classification**
  - Finding a **target distribution** for realistic learning situations
  - Assume **unknown probability distribution** over the input space
  - Hypothesis search with M models and we pick one

- **Statistical Learning Theory**

  \[
  \Pr \left[ \left| E_{in}(g) - E_{out}(g) \right| > \epsilon \right] \leq 2M e^{-2\epsilon^2 N}
  \]

  \[
  \Pr \left[ \left| E_{in}(g) - E_{out}(g) \right| > \epsilon \right] \leq 4m_{\mathcal{H}}(2N) e^{-1/8\epsilon^2 N}
  \]

  **Learning Algorithm** (‘train a system’)
  - ‘learn: get error smaller’
  - ‘generalize well for unseen data’
  - Shift the view to the data and we exchange M with growth function that is indeed depending on N

  **Probability Distribution**
  \[
P \text{ on } X
  \]

  **Hypothesis Set**
  \[
  \mathcal{H} = \{h\}; \ g \in \mathcal{H}
  \]

  **Final Hypothesis**
  \[
  g \approx f
  \]
Outline of the Course

1. Introduction to Machine Learning Fundamentals
2. PRACE and Parallel Computing Basics
3. Unsupervised Clustering and Applications
4. Unsupervised Clustering Challenges & Solutions
5. Supervised Classification and Learning Theory Basics
6. Classification Applications, Challenges, and Solutions
7. Support Vector Machines and Kernel Methods
8. Practicals with SVMs
9. Validation and Regularization Techniques
10. Practicals with Validation and Regularization
11. Parallelization Benefits
12. Cross-Validation Practicals

Day One – beginner
Day Two – moderate
Day Three – expert
Outline

- Remote Sensing Applications
  - Introduction to Application Domain
  - Hyperspectral Datasets
  - Rome Dataset
  - Indian Pines Dataset
  - Explore need for Parallelization

- Classification Challenges & Solutions
  - Selected Key Challenges in Learning from Data
  - Review Practice Experience & Applications
  - Challenge One: Non-Linear Separable Datasets
  - Pocket Algorithm as Extension to PLA
  - Challenge Two: Problem of Overfitting
Remote Sensing Applications
Tsunami December 2004

A devastating tsunami hit many coastal regions on the Indian Ocean

- On of the most devastating natural disasters in recorded history
- 14 countries were hit

How to assess the damage?
Which are the most hit points?
How to plan the Humanitarian aid?
Tsunami December 2004

**Landsat 7**
- 185 Km swath width
- Panchromatic: 15m
- Multispectral (7 bands): 30m
- Thermal band: 60m
- Revisit interval: 16 days

**Quickbird**
- 16.8 Km swath width
- Panchromatic: 65cm
- Multispectral (4 bands): 2.62m
- Revisit time: 1-3.5 days


[4] Landsat 7

Definition and Milestones

**Remote**: without physical contact

**Sensing**: measurement of information

Platforms and Sensors

**Active Sensors:** own source of illumination

- Radar
- Lidar
- Sonar
- Laser Altimeter

**Passive Sensors:** natural light available

- Radiometer
- Imaging Radiometer
- Spectrometer
- Spectroradiometer
Sentinel 2

- Launch dates:
  - Sentinel-2A - 23/06/2015
  - Sentinel-2B – 07/03/2017
- Repeat cycle: 5 days with both satellites
- Multispectral sensor: 13 bands
Application Domains

- Suitable for many applications
- Non invasive method
- Satellite platforms
  - Invaluable view
  - Repetitive and consistent
Application Examples

- Environmental assessment and monitoring
  - Urban growth
  - Hazardous waste
- Global change detection and monitoring
  - Atmospheric ozone depletion
  - Deforestation
  - Global warming


Application Examples

- Military surveillance and reconnaissance
  - Strategic policy
  - Tactical assessment

- Meteorology
  - Atmosphere dynamics
  - Weather prediction
Big Data Analytics

- New analysis challenges
  - 5Vs: Volume, Variety, Velocity, Veracity and Value
- Scalable methods
- Underlying infrastructures

Diagram:

1. DATA ACQUISITION
2. PREPROCESSING
3. PROCESSING
4. INFORMATION ABSTRACTION
5. APPLICATION
Data are complex, noisy and may contain errors:

- Sensors malfunction
- Atmospheric distortion
- Multiple scattering
Data Interpretation

- Once the data is acquired and corrected
- How to interpret the data content and extract information for a specific task?
  - Manual vs Automatic photo-interpertation
Classification of Remote Sensing Images

• Perhaps the most common form of image interpretation

1. Feature transformation
   Spectral or spatial transformation of the input data

2. Training of the classifier
   The classifier is trained on the data (the problem is learned)

3. Labeling of the data
   The classifier performs a decision on all the data
The Value of the Data

- The information content of remote sensing images depends upon various factors
  - The sensor resolution
  - The equipment unreliability
  - The type and amount of noise, etc.
The Spatial Information

QuickBird
Spatial Resolution: 0.6m
The Spectral Information

Improved spectral diversity: hyperspectral imagery
Pixel-Wise Spectral-Spatial Classifiers

Examples of maps obtained by classifying different features

The **spatial features** increases the discrimination of the thematic classes
How to Extract Spatial Features?

- Very High Spatial Resolution images: huge amount of details

- Sub-metric resolution
- Allows for accurate analysis
- Objects with different scales and shapes
Simplification can be performed by attribute filters
They operate only on flat zones according to a criterion
Efficiently implemented on tree representations
Spatial Simplication of the image

Mathematical Morphology connected operators

- Simplification can be performed by **attribute filters**
- They operate only on flat zones according to a criterion
- Efficiently implemented on tree representations

**Flat zone**: set of connected iso-intensity pixels (connected component)
Spatial Simplification of the image

Mathematical Morphology connected operators

- Simplification can be performed by attribute filters
- They operate only on flat zones according to a criterion
- Efficiently implemented on tree representations

Tree representation: each node corresponds to a flat zone
Spatial Simplification of the image

Mathematical Morphology connected operators

- Simplification can be performed by attribute filters
- They operate only on flat zones according to a criterion
- Efficiently implemented on tree representations

Attribute: computed on each node
Spatial Simplification of the image

Mathematical Morphology connected operators

- Simplification can be performed by **attribute filters**
- They operate only on flat zones according to a criterion
- Efficiently implemented on tree representations

Criterion: *binary predicate evaluated on all the nodes*
Spatial Simplification of the image

Mathematical Morphology connected operators

- Simplification can be performed by attribute filters
- They operate only on flat zones according to a criterion
- Efficiently implemented on tree representations

Pruning: removing the nodes that do not fulfill the criterion
Spatial Simplification of the image

Mathematical Morphology connected operators

- Simplification can be performed by **attribute filters**
- They operate only on flat zones according to a criterion
- Efficiently implemented on tree representations

Image simplified: *transform the pruned tree back to an image*
Spatial Simplification of the image

Mathematical Morphology connected operators

- Simplification can be performed by attribute filters
- They operate only on flat zones according to a criterion
- Efficiently implemented on tree representations

Which details should be removed?
It depends on the application

Simplification leads to: removing or attenuating undesired details
Self-Dual Attribute Profiles (SDAPs)

**Representative:** they contain salient structures of the input image

**Non-redundant:** same objects are present only in one or few levels of the SDAP
Example Rome Dataset

- Geographical location: Image of Rome, Italy
  - Multispectral data obtained by Quickbird satellite sensor

- High-resolution (0.6m) panchromatic image
  - Low-resolution (2.4m) multispectral images

(Reasoning for picking SVM: Good classification accuracies on high dimensional datasets, even with a small 'rare' number of training samples)

[Rome Image dataset]

<table>
<thead>
<tr>
<th>Class</th>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buildings</td>
<td>18126</td>
<td>163129</td>
</tr>
<tr>
<td>Blocks</td>
<td>10982</td>
<td>98834</td>
</tr>
<tr>
<td>Roads</td>
<td>16353</td>
<td>147176</td>
</tr>
<tr>
<td>Light Train</td>
<td>1606</td>
<td>14454</td>
</tr>
<tr>
<td>Vegetation</td>
<td>6962</td>
<td>62655</td>
</tr>
<tr>
<td>Trees</td>
<td>9088</td>
<td>81792</td>
</tr>
<tr>
<td>Bare Soil</td>
<td>8127</td>
<td>73144</td>
</tr>
<tr>
<td>Soil</td>
<td>1506</td>
<td>13551</td>
</tr>
<tr>
<td>Tower</td>
<td>4792</td>
<td>43124</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>77542</strong></td>
<td><strong>697859</strong></td>
</tr>
</tbody>
</table>

Lecture 6 – Classification Applications, Challenges, and Solutions
### Inspect and Understanding the Rome Dataset

<table>
<thead>
<tr>
<th>Class</th>
<th>Input</th>
<th>Features</th>
<th>Input</th>
<th>Feature</th>
<th>Input</th>
<th>Features</th>
<th>Input</th>
<th>Features</th>
<th>Input</th>
<th>Features</th>
<th>Input</th>
<th>Features</th>
</tr>
</thead>
</table>

- Each pixel vector is stored as a line with the libSVM format
- E.g.,

```
2 1:0.364706 2:0.360784 3:0.356863 4:0.356863 5:0.349206 6:0.306878 7:0.419355 8:0.453608 9:0.368421 10:1.0 11:1.0 12:0.423529 13:0.403922 14:0.403922 15:0.369919 16:0.320833 17:0.302564 18:0.481481 19:0.483516 20:0.32 21:0.625 22:0.833333 23:0.376471 24:0.376471 25:0.372549 26:0.358566 27:0.318367 28:0.243381 29:0.455446 30:0.4 31:0.319149 32:0.368421 33:0.4 34:0.556863 35:0.549020 36:0.436 37:0.322176 38:0.215962 39:0.151079 40:0.257576 41:0.267857 42:0.266667 43:0.277778 44:0.4375 45:0.360784 46:0.360784 47:0.368627 48:0.368627 49:0.363636 50:0.353846 51:0.347826 52:0.335294 53:0.333333 54:0.978723 55:1
```

[16] G. Cavallaro & M. Riedel et al., 2014
Inspect and Understanding the Rome Dataset

- Data is publicly available in EUDAT B2SHARE tool
Exercises – Explore the Rome Dataset
Example Indian Pines Dataset
NASA – Jet Propulsion Laboratory (JPL)

- Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)
- 224 contiguous spectral channels
- Spectral range of 0.4 μm to 2.5 μm (10 nm spectral resolution)
- Spatial resolution of 20 m
Indian Pine Dataset
June 1992

- Agricultural fields with a variety of crops
- Challenging classification problem
- Similar spectral classes and mixed pixels

[18] Indian Pines dataset
Preprocessing
Corrected by JPL

- 1417 × 617 pixels (~600 MB)
- 200 bands (20 discarded, with low SNR)
- 58 classes (6 discarded, with ≤ 100 samples)

---

<table>
<thead>
<tr>
<th>number</th>
<th>Class</th>
<th>Number of samples</th>
<th></th>
<th>number</th>
<th>Class</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>name</td>
<td>training</td>
<td>test</td>
<td></td>
<td>name</td>
<td>training</td>
</tr>
<tr>
<td>1</td>
<td>Buildings</td>
<td>1720</td>
<td>1275</td>
<td>27</td>
<td>Pasture</td>
<td>1039</td>
</tr>
<tr>
<td>2</td>
<td>Corn</td>
<td>1776</td>
<td>1005</td>
<td>28</td>
<td>good</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Corn?</td>
<td>16</td>
<td>142</td>
<td>29</td>
<td>Soybeans?</td>
<td>939</td>
</tr>
<tr>
<td>4</td>
<td>Corn-FW</td>
<td>51</td>
<td>463</td>
<td>30</td>
<td>Soybeans?</td>
<td>89</td>
</tr>
<tr>
<td>5</td>
<td>Corn-NS</td>
<td>236</td>
<td>2120</td>
<td>31</td>
<td>Soybeans-NS</td>
<td>111</td>
</tr>
<tr>
<td>6</td>
<td>Corn-CleanTill</td>
<td>1240</td>
<td>11664</td>
<td>32</td>
<td>Soybeans-CleanTill</td>
<td>597</td>
</tr>
<tr>
<td>7</td>
<td>Corn-CleanTill-EW</td>
<td>2049</td>
<td>23837</td>
<td>33</td>
<td>Soybeans-CleanTill</td>
<td>273</td>
</tr>
<tr>
<td>8</td>
<td>Corn-CleanTill-NS</td>
<td>3968</td>
<td>35710</td>
<td>34</td>
<td>Soybeans-CleanTill-NS</td>
<td>1180</td>
</tr>
<tr>
<td>9</td>
<td>Corn-CleanTill-NS-Irrigated</td>
<td>39</td>
<td>720</td>
<td>35</td>
<td>Soybeans-CleanTill-NS-Irrigated</td>
<td>1029</td>
</tr>
<tr>
<td>10</td>
<td>Corn-CleanTilled-NS-Irrigated</td>
<td>173</td>
<td>1555</td>
<td>36</td>
<td>Soybeans-CleanTill-Damaged</td>
<td>224</td>
</tr>
<tr>
<td>11</td>
<td>Corn-MixTill</td>
<td>108</td>
<td>944</td>
<td>37</td>
<td>Soybeans-CleanTill-Weedy</td>
<td>54</td>
</tr>
<tr>
<td>12</td>
<td>Corn-MixTill-EW</td>
<td>560</td>
<td>5065</td>
<td>38</td>
<td>Soybeans-MixTill-Drilled</td>
<td>151</td>
</tr>
<tr>
<td>13</td>
<td>Corn-MixTill-NS</td>
<td>380</td>
<td>7976</td>
<td>39</td>
<td>Soybeans-MixTill-Drilled</td>
<td>267</td>
</tr>
<tr>
<td>14</td>
<td>Corn-NoTill</td>
<td>438</td>
<td>3943</td>
<td>40</td>
<td>Soybeans-NoTill</td>
<td>183</td>
</tr>
<tr>
<td>15</td>
<td>Corn-NoTill-EW</td>
<td>121</td>
<td>1085</td>
<td>41</td>
<td>Soybeans-NoTill-Drilled</td>
<td>810</td>
</tr>
<tr>
<td>16</td>
<td>Corn-NoTill-NS</td>
<td>260</td>
<td>5116</td>
<td>42</td>
<td>Soybeans-NoTill-NS-Irrigated</td>
<td>495</td>
</tr>
<tr>
<td>17</td>
<td>Peas</td>
<td>11</td>
<td>103</td>
<td>43</td>
<td>Soybeans-NoTill</td>
<td>216</td>
</tr>
<tr>
<td>18</td>
<td>Grass</td>
<td>115</td>
<td>1032</td>
<td>44</td>
<td>Soybeans-NoTill-Drilled</td>
<td>253</td>
</tr>
<tr>
<td>19</td>
<td>Grass-Trees</td>
<td>233</td>
<td>2098</td>
<td>45</td>
<td>Soybeans-NoTill-NS-Irrigated</td>
<td>91</td>
</tr>
<tr>
<td>20</td>
<td>Hay</td>
<td>313</td>
<td>1015</td>
<td>46</td>
<td>Soybeans-NoTill-Drilled</td>
<td>873</td>
</tr>
<tr>
<td>21</td>
<td>Hay?</td>
<td>24</td>
<td>1966</td>
<td>47</td>
<td>Swampy Area</td>
<td>58</td>
</tr>
<tr>
<td>22</td>
<td>Hay-Alalfa</td>
<td>222</td>
<td>2052</td>
<td>48</td>
<td>River</td>
<td>311</td>
</tr>
<tr>
<td>23</td>
<td>Lake</td>
<td>2</td>
<td>202</td>
<td>49</td>
<td>Trees</td>
<td>58</td>
</tr>
<tr>
<td>24</td>
<td>NotCropped</td>
<td>194</td>
<td>1746</td>
<td>50</td>
<td>Wheat</td>
<td>495</td>
</tr>
<tr>
<td>25</td>
<td>Oats</td>
<td>174</td>
<td>1585</td>
<td>51</td>
<td>Woods</td>
<td>6356</td>
</tr>
<tr>
<td>26</td>
<td>Oats?</td>
<td>34</td>
<td>301</td>
<td>52</td>
<td>Woods?</td>
<td>14</td>
</tr>
</tbody>
</table>

Experimental Setup

Two Cases

Feature Enhancement & Selection
- Kernel Principle Component Analysis (KPCA)
- Extended Self-Dual Attribute Profile (ESDAP)
- Nonparametric weighted feature extraction (NWFE)

Inspect and Understanding the Indian Pines Dataset

- Dataset raw (1)
  - Class + Original Spectral Bands
    - 200 spectral bands
    - 48 1:0.365 2:0.361 3:0.356 ...
    - 209:0.333 220:0.978
    - libSVM

- Dataset processed (2)
  - Class + Transformed Features
    - 30 image features
    - 48 1:0.245 2:0.360 3:0.326 ...
    - 29:0.241 30:0.878
    - libSVM
Available Datasets – Training Data Example

- **Indian Pines Dataset Processed – Training**
  - Indian_produced_training.el
  - LibSVM data format: class feature1:value1 feature2:value2

---

Lecture 6 – Classification Applications, Challenges, and Solutions
Publicly Available Datasets – Location

- **Indian Pines Dataset Raw and Processed**

  ![Indian Pines Dataset Raw and Processed](image_url)

  **Abstract:**
  1. Indian raw: 1417x614x200 (training 10% and test) 2. Indian processed: 1417x614x30 (training 10% and test)

  **PID:** 11304/7e8eeec8e-3d61-11e4-ac7e-860a000653d1

  **Files:**
  - `indian_processed_test.el` 105.69MB
  - `indian_processed_training.el` 117.3MB
  - `indian_raw_test.el` 74713MB
  - `indian_raw_training.el` 8301311MB

  ![Basic metadata](image_url)

  - **Open Access:** True
  - **License:**
  - **Contact Email:**
  - **Publication Date:** 2015-02-04
  - **Contributors:**
  - **Resource Type:**
  - **Alternate identifiers:** 172
  - **Type:** B2SHARE.V1_ID
  - **Publisher:** https://b2share.eudat.eu
  - **Language:** en
Exercises – Explore Indian Pines Dataset
[Video] Remote Sensing

Classification Challenges & Solutions
Key Challenges: Why is it not so easy in practice?

- **Scalability**
  - Gigabytes, Terabytes, and Petabytes datasets that fit not into memory
  - E.g. algorithms become necessary with out-of-core/CPU strategies

- **High Dimensionality**
  - Datasets with hundreds or thousand attributes become available
  - E.g. bioinformatics with gene expression data with thousand of features

- **Heterogenous and Complex Data**
  - More complex data objects emerge and unstructured data sets
  - E.g. Earth observation time-series data across the globe

- **Data Ownership and Distribution**
  - Distributed datasets are common (e.g. security and transfer challenges)
Challenge One - Non-linearly Separable Data in Practice

(4) Modelling Phase

\[(x_1, y_1), \ldots, (x_N, y_N)\]
(resampled, again
\(N = 100\) samples)

(lessons learned from practice: requires
soft-thresholds to allow for some errors
being overall better for new data)

(linear decision boundary)

(non-linear decision boundary)

(lessons learned from practice: requires
non-linear decision boundaries)
Solution Tools: Linear Perceptron Hypothesis Set & Pocket

- **Unknown Target Distribution**
  - Target function $f : X \rightarrow Y$ plus noise $P(y|x)$

- **Probability Distribution**
  - $P$ on $X$

- **Training Examples**
  - $(x_1, y_1), \ldots, (x_N, y_N)$

- **Error Measure**
  - $e(x)$

- **Learning Algorithm**
  - ‘train a system’
  - (Pocket Algorithm)

- **Final Hypothesis**
  - $g \approx f$

- **Hypothesis Set**
  - $\mathcal{H} = \{h\}; \; g \in \mathcal{H}$

- **Elements we not exactly (need to) know**
  - ‘constants’ in learning

- **Elements we must and/or should have and that might raise huge demands for storage**

- **Elements that we derive from our skillset and that can be computationally intensive**

- **Elements that we derive from our skillset**
Smart Adhancement of PLA – Pocket Algorithm

- When: If we believe there is a **linear pattern** to be detected
  - No assumption: can be non-linearly separable data

- Basis is still the PLA
- Idea: Put the best solution so far ‘in a pocket’
- Best means: Error measure checks per iterations
- Works with non-linearly separable data
- Needs **fixed iterations number** (otherwise no convergence of algorithm)
Challenge Two – Problem of Overfitting

- Overfitting refers to fit the data too well – more than is warranted – thus may misguide the learning
- Overfitting is not just ‘bad generalization‘ - e.g. the VC dimension covers noiseless & noise targets
- Theory of Regularization are approaches against overfitting and prevent it using different methods

- Key problem: **noise in the target function leads to overfitting**
  - Effect: ‘noisy target function‘ and its noise misguides the fit in learning
  - There is always ‘some noise‘ in the data
  - Consequence: poor target function (‘distribution’) approximation

- Example: Target functions is **second order polynomial** (i.e. parabola)
  - Using a higher-order polynomial fit
  - Perfect fit: low $E_{in}(g)$, but large $E_{out}(g)$

- Overfitting refers to fit the data too well – more than is warranted – thus may misguide the learning
- Overfitting is not just ‘bad generalization‘ - e.g. the VC dimension covers noiseless & noise targets
- Theory of Regularization are approaches against overfitting and prevent it using different methods

- Key problem: **noise in the target function leads to overfitting**
  - Effect: ‘noisy target function‘ and its noise misguides the fit in learning
  - There is always ‘some noise‘ in the data
  - Consequence: poor target function (‘distribution’) approximation

- Example: Target functions is **second order polynomial** (i.e. parabola)
  - Using a higher-order polynomial fit
  - Perfect fit: low $E_{in}(g)$, but large $E_{out}(g)$

- Overfitting refers to fit the data too well – more than is warranted – thus may misguide the learning
- Overfitting is not just ‘bad generalization‘ - e.g. the VC dimension covers noiseless & noise targets
- Theory of Regularization are approaches against overfitting and prevent it using different methods

- Key problem: **noise in the target function leads to overfitting**
  - Effect: ‘noisy target function‘ and its noise misguides the fit in learning
  - There is always ‘some noise‘ in the data
  - Consequence: poor target function (‘distribution’) approximation

- Example: Target functions is **second order polynomial** (i.e. parabola)
  - Using a higher-order polynomial fit
  - Perfect fit: low $E_{in}(g)$, but large $E_{out}(g)$
Problem of Overfitting – Clarifying Terms

- A good model must have low training error ($E_{in}$) and low generalization error ($E_{out}$).
- Model overfitting is if a model fits the data too well ($E_{in}$) with a poorer generalization error ($E_{out}$) than another model with a higher training error ($E_{in}$).

### Overfitting & Errors
- $E_{in}(g)$ goes down
- $E_{out}(g)$ goes up
- ‘Bad generalization area’ ends
  - Good to reduce $E_{in}(g)$
- ‘Overfitting area’ starts
  - Reducing $E_{in}(g)$ does not help
  - Reason ‘fitting the noise’

### The two general approaches to prevent overfitting are (1) regularization and (2) validation

Lecture 9 provides details on validation to be considered as another method against overfitting.
[Video] Overfitting Challenge

So, let me state that again.

Degrees of freedom

[22] YouTube Video, ‘Overfitting’
Lecture Bibliography (1)

- [1] The 2004 Indian Ocean Tsunami
  Online: https://www.thoughtco.com/the-2004-indian-ocean-tsunami-195145
- [2] Five years after Indian Ocean tsunami, affected nations rebuilding better – UN
  Online: https://earthobservatory.nasa.gov/IOTD/view.php?id=5145
- [4] Landsat 7: the seventh satellite of the Landsat program
  Online: https://landsat.gsfc.nasa.gov/landsat-7/
  Online: https://www.satimagingcorp.com/satellite-sensors/quickbird/
  Online: https://www.pinterest.com/pin/291608144593636608/
- [7] Aerial photographs of Schlosshotel Kronberg, near Frankfurt, Germany
  Online: https://en.wikipedia.org/wiki/Pigeon_photography
- [8] First aerial photography taken from an airplane, 1909, near Rome, Italy
  Online: http://professionalaerialphotographers.com/content.aspx?page_id=22&club_id=808138&module_id=158950
  Online: https://www.oneonta.edu/faculty/baumanpr/geosat2/RS%20History%20II/RS-History-Part-2.html
- [10] Sentinel-2 global coverage
  Online: https://www.youtube.com/watch?v=upXdz6Qvl3c
Online: https://manilabydaniellaandisabel.weebly.com/location-and-characteristics.html

Online: https://www.satimagingcorp.com/gallery/more-imagery/aster/aster-deforestation-bolivia/

Online: https://commons.wikimedia.org/wiki/File:Temporary_Armor_Storage_Area,_Managua,_Nicaragua,_11_November_1981_%E2%80%93_CIA_IMINT.jpg

Online: http://www.noaa.gov

[15] Rome Image dataset
Online: https://b2share.eudat.eu/records/daf6c389e54340b4b1416cf874251e77


[17] Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)
Online: https://www.jpl.nasa.gov/missions/airborne-visible-infrared-imaging-spectrometer-aviris/

[18] Indian Pines dataset: 220 Band AVIRIS Hyperspectral Image
Online: https://purr.purdue.edu/publications/1947/1


[20] Indian Pines Raw and Processed
Online: http://hdl.handle.net/11304/9ec5eac8-61b4-4617-ae1c-1f8c8cd3cd74
Lecture Bibliography (3)

- [22] YouTube Video, ‘Overfitting’, Online: https://www.youtube.com/watch?v=mfzHchd5La8