Reflections on Data Science for Finance (DSfin)

Louiqa Raschid
Smith School of Business
Computer Science and UMIACS

Thanks to NSF grants CNS1305368 and IIP1464644 and OFR / NIST award 70NANB15H194.
DSfin Challenges

- DSfin supply chains and technical challenges:
  - Lack of ontologies and standards including identifiers!
  - Absence of OPEN datasets.
  - Time-series longitudinal datasets versus graphs or semi-structured datasets.

- Unique to the DSfin domain:
  - Structural: Financial markets operate in silos.
  - Cultural: Awash in $$$; information asymmetry; lack of benchmarks; etc.
(Toxic) Financial Entities and Impact on resMBS Securities

Louiqa Raschid
Zheng Xu and and Elena Zotkina

Thanks to NSF grants CNS1305368 and IIP1464644
Outline

• Motivation
• Supply chain for residential mortgage backed securities.
• Latent “Financial Community” model using LDA based probabilistic model.
• Examples of “toxic” financial entities / communities / securities.
• Impact of upstream toxicity along the supply chain on downstream securities: ratings and performance.
• Ongoing work ...
Mortgages

Mortgage Pool

Mortgage Backed Security
a.k.a. “Collateralized Debt Obligation”
(divided into “tranches” and sold at different levels of risk)

Risk

Unsecured 12%
Mezzanine 9%
Sr. Secured 7%

Expected Return

a.k.a. “Derivatives”
(broken up into groups and ‘derived’ from the original pool)
Mortgages

Mortgage Pool

Mortgage Backed Security
a.k.a. “Collateralized Debt Obligation”
(divided into “tranches” and sold at different levels of risk)

Risk

Unsecured → 12%
Mezzanine → 9%
Sr. Secured → 7%

Expected Return

a.k.a. “Derivatives” (broken up into groups and ‘derived’ from the original pool)
Constructing the upstream resMBS supply chain

<table>
<thead>
<tr>
<th>Role</th>
<th>Entity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title of Series</td>
<td>Wachovia Mortgage Loan Trust, LLC Mortgage Pass-Through Certificates, Series 2007-A.</td>
</tr>
<tr>
<td>Depositor</td>
<td>Wachovia Mortgage Loan Trust, LLC.</td>
</tr>
<tr>
<td>Issuing entity</td>
<td>Wachovia Mortgage Loan Trust, LLC Series 2007-A Trust.</td>
</tr>
<tr>
<td>Seller and Sponsor</td>
<td>Wachovia Bank, National Association.</td>
</tr>
<tr>
<td>Originators</td>
<td>National City Mortgage, Fifth Third Mortgage Company, SunTrust Mortgage, Inc. and Wells Fargo Bank, N.A.</td>
</tr>
<tr>
<td>Servicers</td>
<td>National City Mortgage Co., or National City, Fifth Third Mortgage Company, SunTrust Mortgage, Inc. and Wells Fargo Bank, N.A.</td>
</tr>
<tr>
<td>Trustee</td>
<td>HSBC Bank USA, National Association.</td>
</tr>
</tbody>
</table>
Securities (2004) from a “non-toxic” community
Securities issued in 2004 associated with a “toxic” financial community “FIRST HORIZON” ...
DSfin Challenges

● DSfin supply chains and technical challenges:
  ○ Lack of ontologies and standards including identifiers!
  ○ Absence of OPEN datasets.
  ○ Time-series longitudinal datasets versus graphs or semi-structured datasets.

● Unique to the DSfin domain:
  ○ Structural: Financial markets operate in silos.
  ○ Cultural: Awash in $$$; information asymmetry; lack of benchmarks; etc.