Big Data Analytics

Lucas Rego Drumond

Information Systems and Machine Learning Lab (ISMLL)
Institute of Computer Science
University of Hildesheim, Germany

Big Data Analytics

Outline

1. GraphLab Application Deployment

2. Relational Classification Example

Outline

Steps

1. Install GraphLab into a specific directory
   - Example: `/home/user/graphlab`
2. Create a directory for your application under `/home/user/graphlab/apps`
3. Create a CMakeLists.txt file into your application directory
4. Add the source files for your program under your application directory
5. Run `./configure` under `/home/user/graphlab`
6. Go to `/home/user/graphlab/release/apps/your_application` and type `make`

CMakeLists.txt

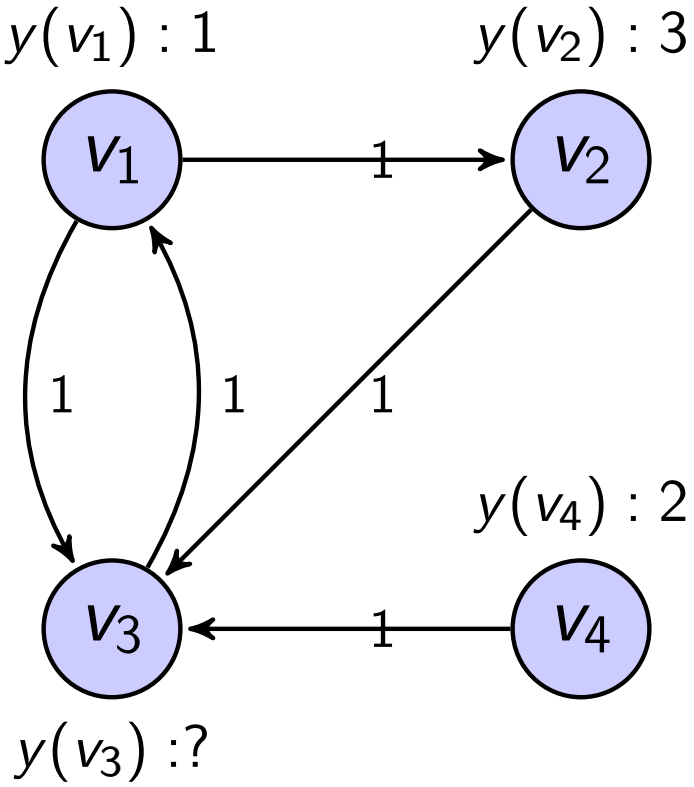
```
project(MyProjectName)

add_graphlab_executable(executable_name implementation.cpp)
```

Hello World

```
#include <graphlab.hpp>

int main(int argc, char** argv) {

  graphlab::mpi_tools::init(argc, argv);
  graphlab::distributed_control dc;

  dc.cout() << "Hello World!\n";

  graphlab::mpi_tools::finalize();

}
```

Outline

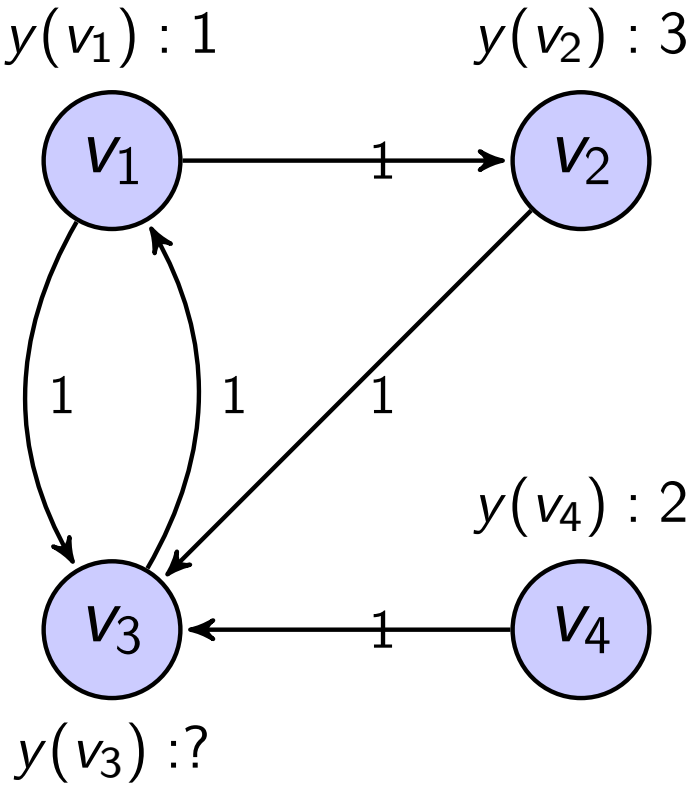
Relational Classification



Given a graph $G := (V, E)$ and a set of labels L

- Some nodes have labels $y : V \rightarrow L$
- Edges v, u have weights $w_{v,u}$
- Task: estimate a function $\hat{y} : V \rightarrow L$

Weighted voted Relational Neighbor



Probability that a vertex $v \in V$ has label $c \in L$

$$P(c|v) = \frac{1}{Z_v} \sum_{u \in \{u | u \in \mathcal{N}_v \wedge y(u) = c\}} w_{(u,v)}$$

Where:

$$Z_v = \sum_{u \in \mathcal{N}_v} w_{(u,v)}$$

- $\mathcal{N}_v$ denotes the neighbors of v

$$\hat{y}(v) := \arg\max_{c \in L} P(c|v)$$

wvRN Vertex Program:

1: **procedure** wvRNGather
input: vertex v, scope $\mathcal{S}_v$, ingoing edge $(u \rightarrow v)$

2: **return** $(w_{(u,v)}, y(u))$

3: **end procedure**

1: **procedure** wvRNApply
input: vertex v, scope $\mathcal{S}_v$, gather result $(Z_v, \left(\sum_{\{u|u \in \mathcal{N}_v \wedge y(u)=c\}} w_{u,v}\right)_{c \in L})$,

2: $\hat{y}(v) := \arg\max_{c \in L} \left(\sum_{\{u|u \in \mathcal{N}_v \wedge y(u)=c\}} w_{u,v}\right)$

3: **end procedure**

wvRN Code

- Code and toy data:
  - `http://www.ismll.uni-hildesheim.de/lehre/bd-14s/script/gl_ex/wvRN_example.zip`