SEMI-SUPERVISED STANCE DETECTION IN TWEETS BASED ON SENTIMENT RULES

Marcelo Dias and Karin Becker
Instituto de Informática – UFRGS – Porto Alegre - Brazil
marcelo.dias@inf.ufrgs.br and karin.becker@inf.ufrgs.br

Introduction

- Opinion Analysis
  - Detect sentiment polarity (negative or positive)
  - Target (often mentioned in the text)

- Stance Detection
  - Detect Stance (against or favor)
  - Towards a given target (main target vs indirect targets)
  - In favor stance can be expressed through positive/negative sentiments (and vice-versa)

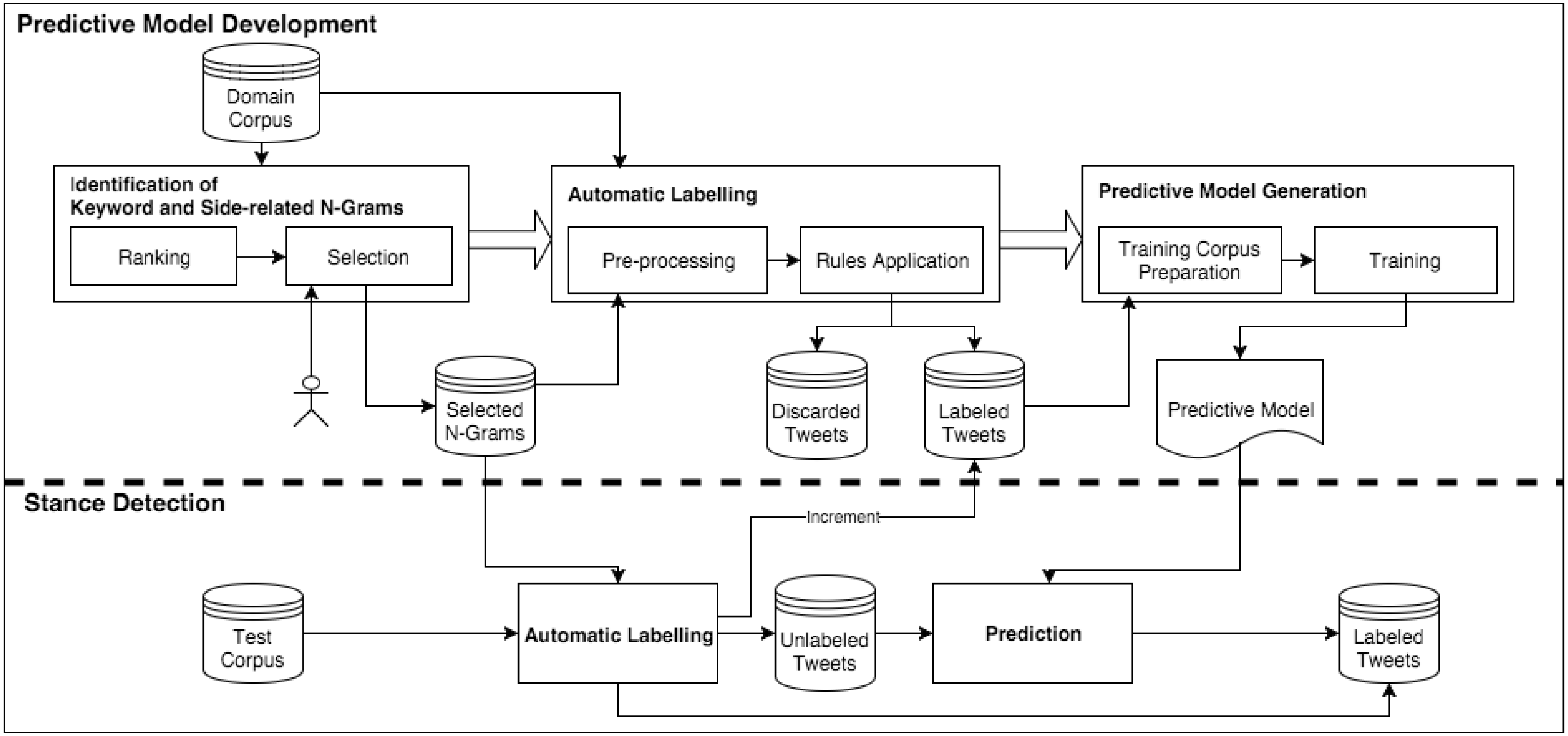
Introduction

- Related Work
  - Structured text or discussion threads (congress vote, on-line debate, ...)
    - wider textual context to interpret content
    - [Thomas et al. 2006] [Anand et al. 2011]
    [Somasundaran and Wiebe 2009]
  - Tweets: short text and poorly written content
    - rely more on inferences from static/dynamic properties of the platform
      - [Rajadesingan and Liu 2014]
    - Less focus on properties extracted from textual contents only
  - Most works adopt supervised methods
  - Often address a binary problem (Favor/Against)

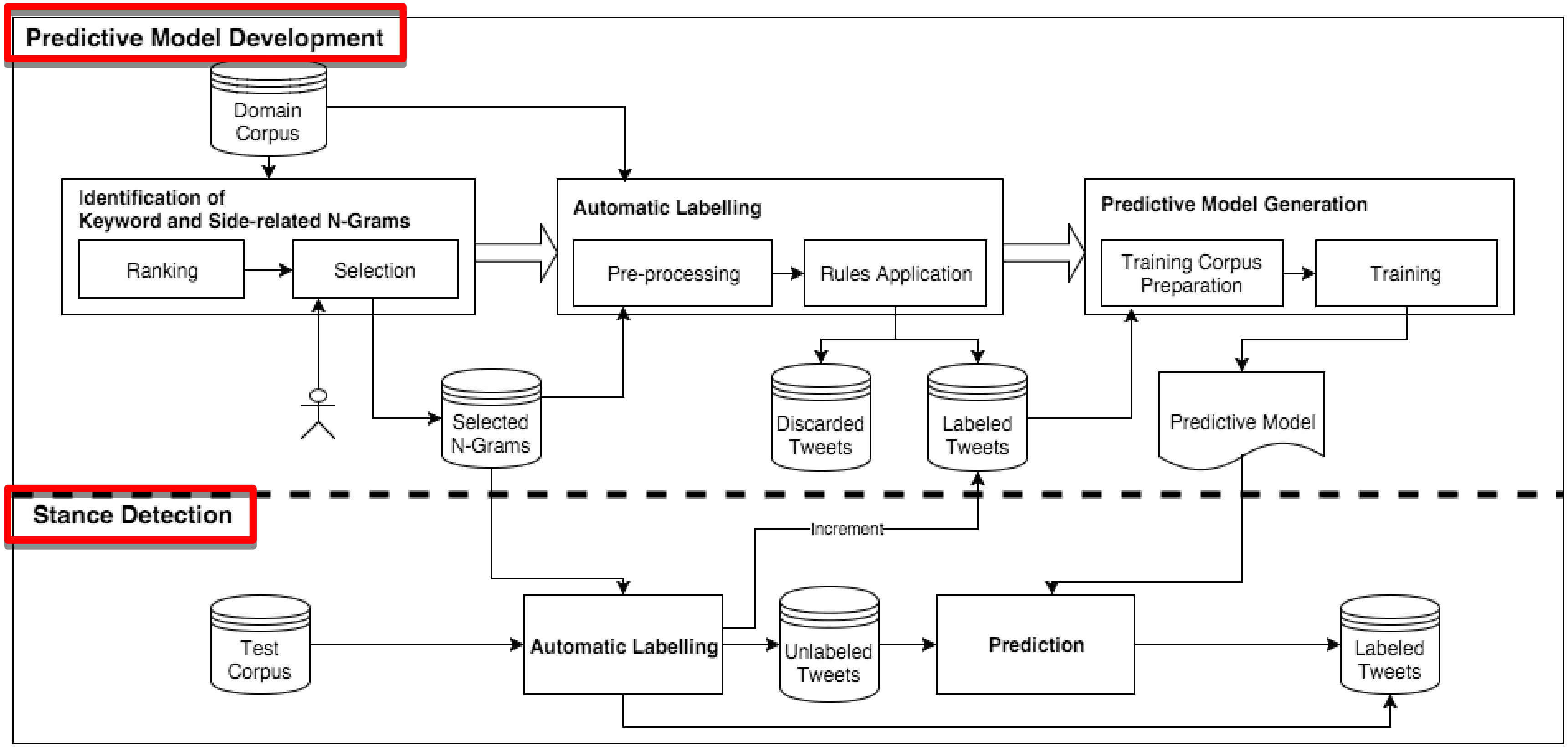
Goal

- Stance Detection based only on tweets textual content
  - Rule-based, Semi-supervised method
  - 3 classes problem (Favor, Against and None)
  - Improvements on our early work
    - Third place in SemEval 2016 Task 6-B (unsupervised, Trump Target)
  - Evaluate generality using several distinct domains
    - SemEval 2016 Task 6-A Targets (supervised)

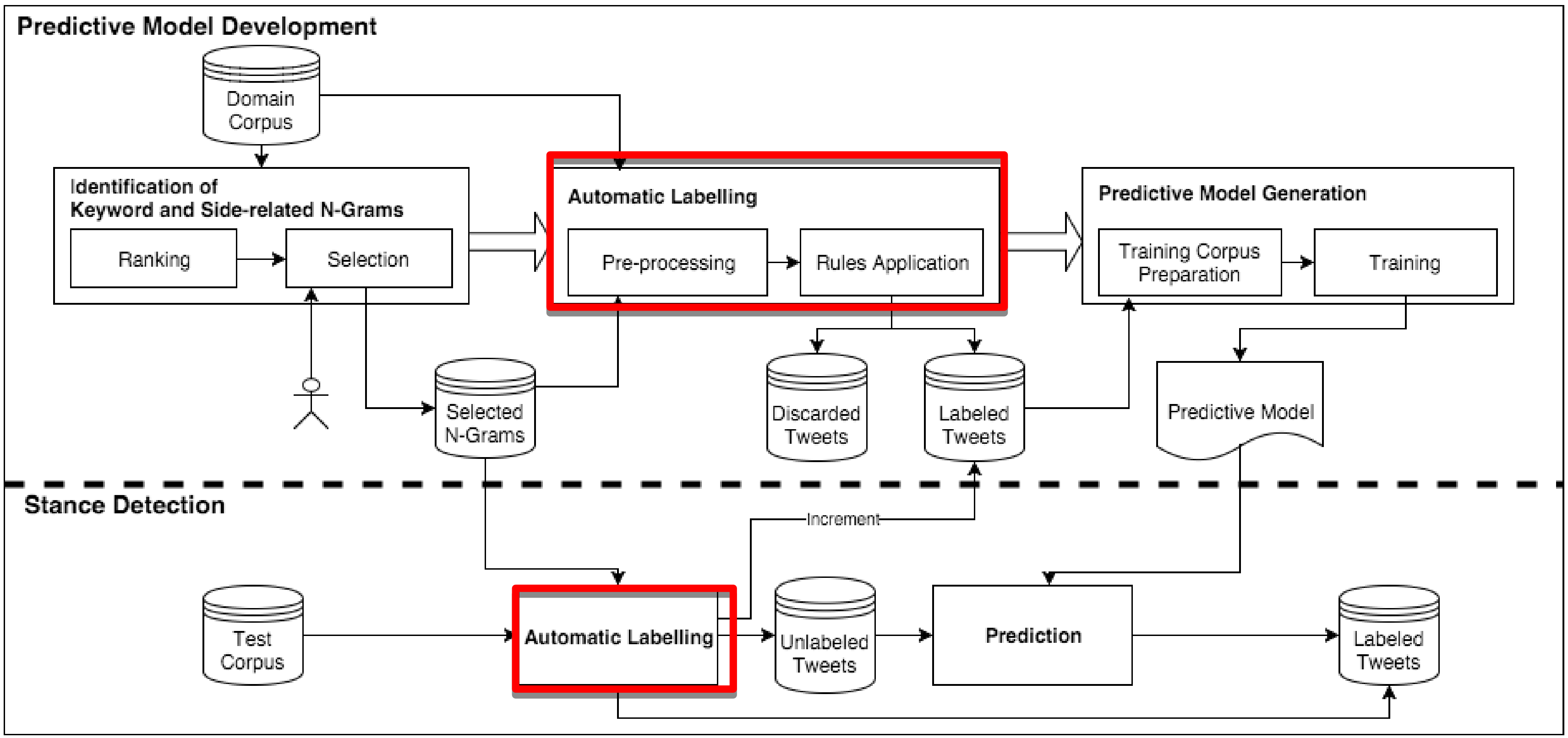
Process Overview



Process Overview



Process Overview: automatic labeling



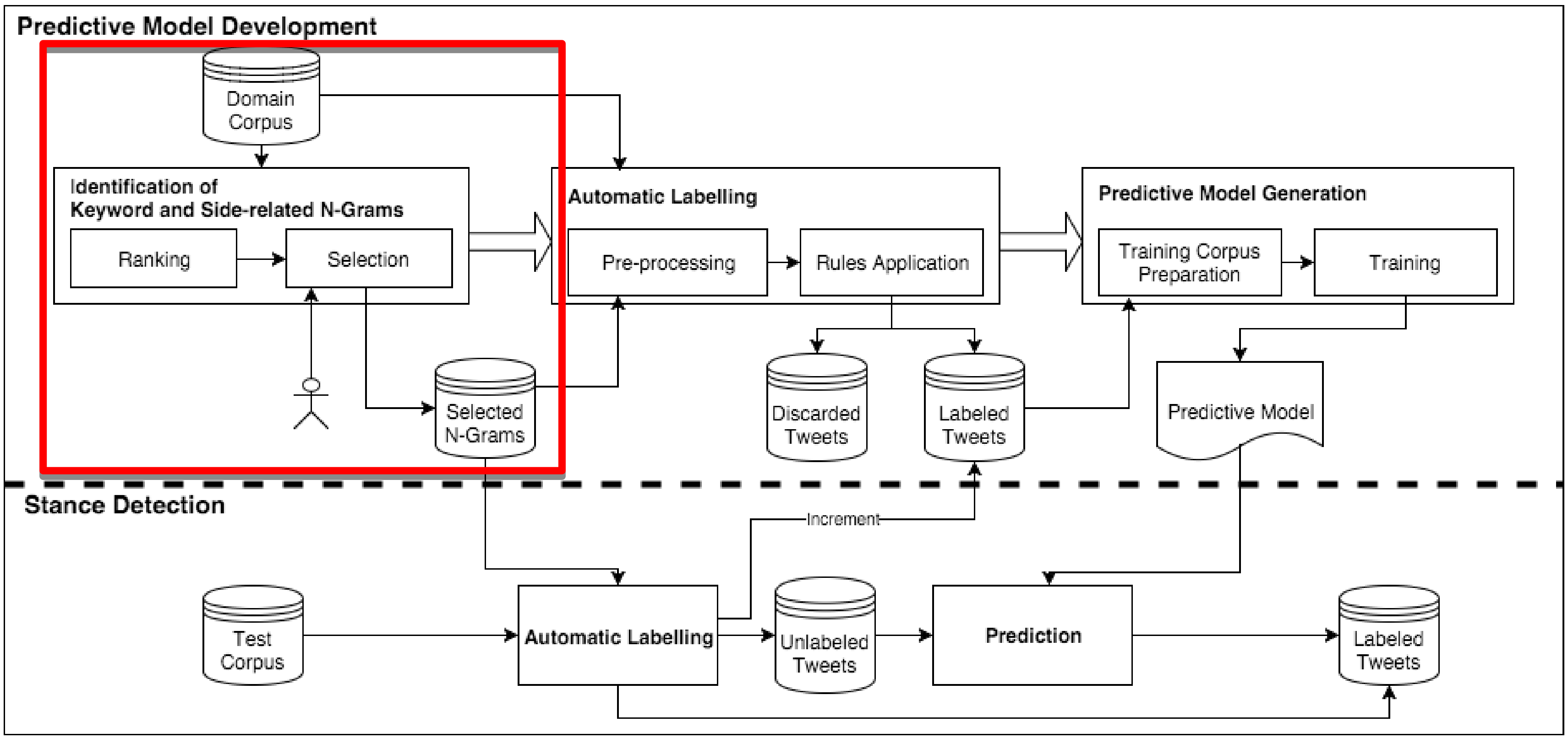
Key and Target N-grams

- Key n-grams: terms/phrases that denote a stance
- Target n-grams: identify a target directly or indirectly related to main target
  - combined with polarity to denote a stance
- May be Favor or Against

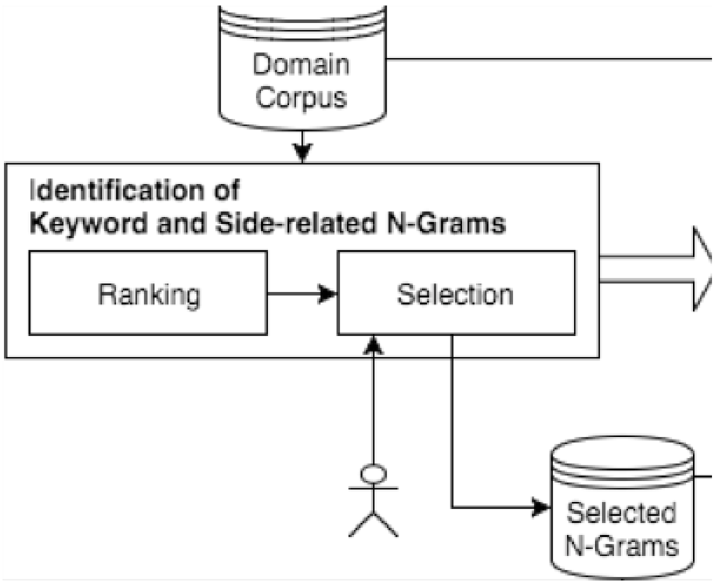
Main target: Hillary Clinton

N-GRAMS	FAVOR	AGAINST
KEY	ReadyForHillary, Hillary2016	StopHillary, MakeAmericaGreatAgain
TARGET	Hillary, Democrats	Trump, Republicans

Key and Target N-grams Identification



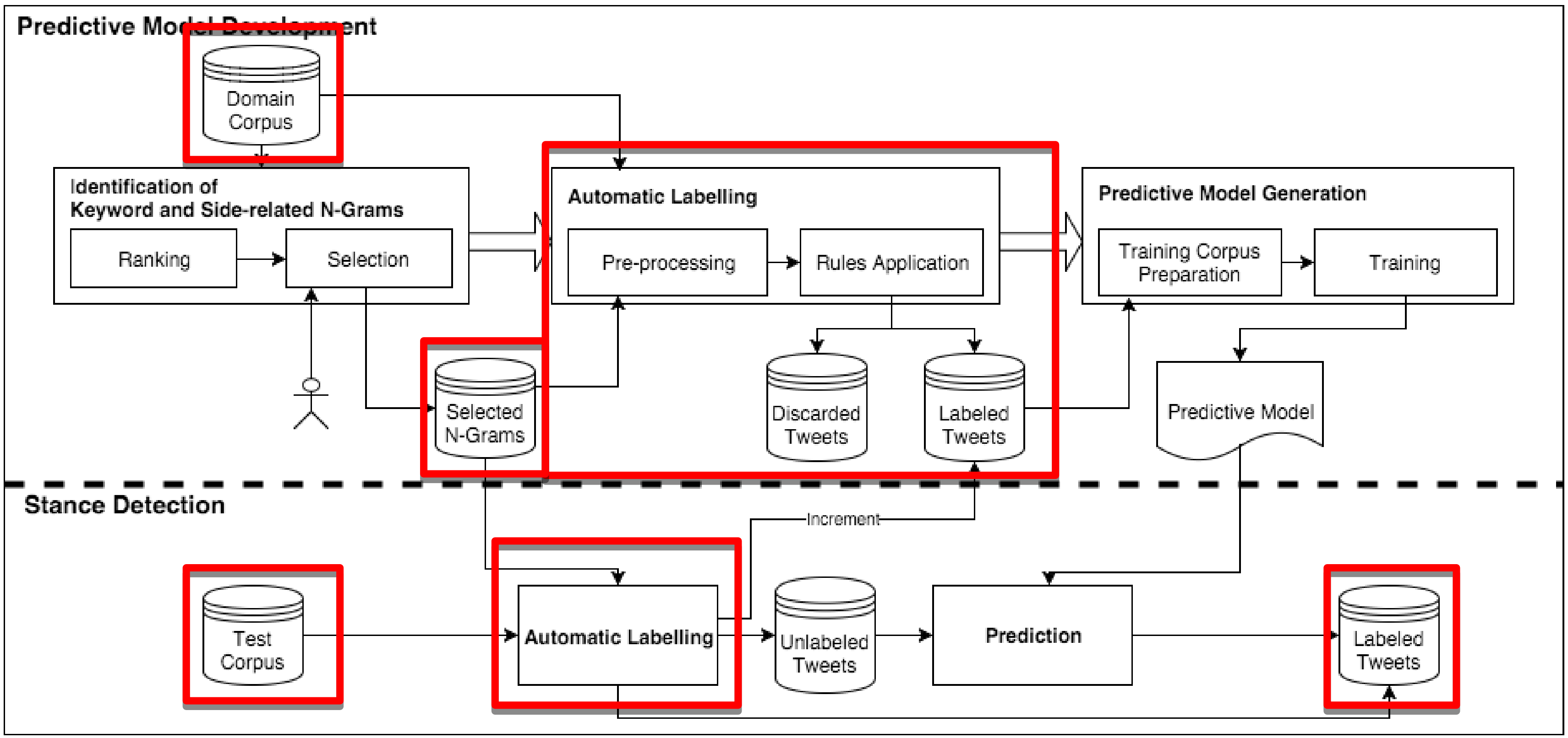
Key and Target N-grams Identification



- Input: domain corpus
- Current selection
  - N-Gram frequency ranking
  - Manual selection of top frequent n-grams
- Output: selected Key and Target n-grams

- Currently evaluating automatic n-grams selection methods

Process Overview: Automatic Labeling



Rules x Stance

Rule	Stance
1 - KEY-FAVOR Presence of a favor keyword n-gram with no against keyword n-gram	FAVOR
2 - KEY-AGAINST Presence of an against keyword n-gram with no favor keyword n-gram	AGAINST
3 - FAVOR-POSITIVE Presence of a favor-related side target with no against-related side target and positive tweet polarity	FAVOR
4 - FAVOR-NEGATIVE Presence of a favor-related side target with no against-related side target and negative tweet polarity	AGAINST
5 - AGAINST-POSITIVE Presence of a against side related target with no favor-related side target and positive tweet polarity	AGAINST
6 - AGAINST-NEGATIVE Presence of an against-related side target with no favor-related side target and negative tweet polarity	FAVOR
7 - NONE Absense of keyword n-grams, side related n-grams and hashtags	NONE
Other cases	DISCARD TWEET

FEATURES
Presence of at least one Favor/Against Key N-grams
Presence of at least one Favor/Against Target N-grams
Presence of at least one hashtag
Tweet Polarity

Rules x Stance

Rule	Stance
1 - KEY-FAVOR Presence of a favor keyword n-gram with no against keyword n-gram	FAVOR
2 - KEY-AGAINST Presence of an against keyword n-gram with no favor keyword n-gram	AGAINST
3 - FAVOR-POSITIVE Presence of a favor-related side target with no against-related side target and positive tweet polarity	FAVOR
4 - FAVOR-NEGATIVE Presence of a favor-related side target with no against-related side target and negative tweet polarity	AGAINST
5 - AGAINST-POSITIVE Presence of a against side related target with no favor-related side target and positive tweet polarity	AGAINST
6 - AGAINST-NEGATIVE Presence of an against-related side target with no favor-related side target and negative tweet polarity	FAVOR
7 - NONE Absense of keyword n-grams, side related n-grams and hashtags	NONE
Other cases	DISCARD TWEET

FEATURES
Presence of at least one Favor/Against Key N-grams
Presence of at least one Favor/Against Target N-grams
Presence of at least one hashtag
Tweet Polarity

Rules x Stance

Rule	Stance
1 - KEY-FAVOR Presence of a favor keyword n-gram with no against keyword n-gram	FAVOR
2 - KEY-AGAINST Presence of an against keyword n-gram with no favor keyword n-gram	AGAINST
3 - FAVOR-POSITIVE Presence of a favor-related side target with no against-related side target and positive tweet polarity	FAVOR
4 - FAVOR-NEGATIVE Presence of a favor-related side target with no against-related side target and negative tweet polarity	AGAINST
5 - AGAINST-POSITIVE Presence of a against side related target with no favor-related side target and positive tweet polarity	AGAINST
6 - AGAINST-NEGATIVE Presence of an against-related side target with no favor-related side target and negative tweet polarity	FAVOR
7 - NONE Absense of keyword n-grams, side related n-grams and hashtags	NONE
Other cases	DISCARD TWEET

FEATURES
Presence of at least one Favor/Against Key N-grams
Presence of at least one Favor/Against Target N-grams
Presence of at least one hashtag
Tweet Polarity

Rules x Stance

Rule	Stance
1 - KEY-FAVOR Presence of a favor keyword n-gram with no against keyword n-gram	FAVOR
2 - KEY-AGAINST Presence of an against keyword n-gram with no favor keyword n-gram	AGAINST
3 - FAVOR-POSITIVE Presence of a favor-related side target with no against-related side target and positive tweet polarity	FAVOR
4 - FAVOR-NEGATIVE Presence of a favor-related side target with no against-related side target and negative tweet polarity	AGAINST
5 - AGAINST-POSITIVE Presence of a against side related target with no favor-related side target and positive tweet polarity	AGAINST
6 - AGAINST-NEGATIVE Presence of an against-related side target with no favor-related side target and negative tweet polarity	FAVOR
7 - NONE Absense of keyword n-grams, side related n-grams and hashtags	NONE
Other cases	DISCARD TWEET

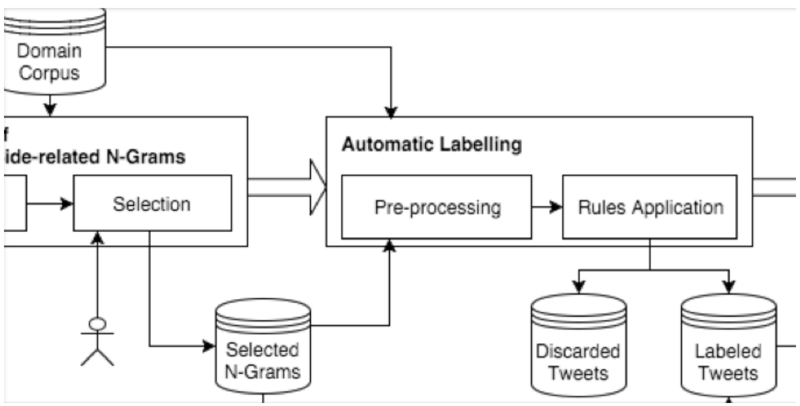
FEATURES
Presence of at least one Favor/Against Key N-grams
Presence of at least one Favor/Against Target N-grams
Presence of at least one hashtag
Tweet Polarity

Rules x Stance

Rule	Stance
1 - KEY-FAVOR Presence of a favor keyword n-gram with no against keyword n-gram	FAVOR
2 - KEY-AGAINST Presence of an against keyword n-gram with no favor keyword n-gram	AGAINST
3 - FAVOR-POSITIVE Presence of a favor-related side target with no against-related side target and positive tweet polarity	FAVOR
4 - FAVOR-NEGATIVE Presence of a favor-related side target with no against-related side target and negative tweet polarity	AGAINST
5 - AGAINST-POSITIVE Presence of a against side related target with no favor-related side target and positive tweet polarity	AGAINST
6 - AGAINST-NEGATIVE Presence of an against-related side target with no favor-related side target and negative tweet polarity	FAVOR
7 - NONE Absense of keyword n-grams, side related n-grams and hashtags	NONE
Other cases	DISCARD TWEET

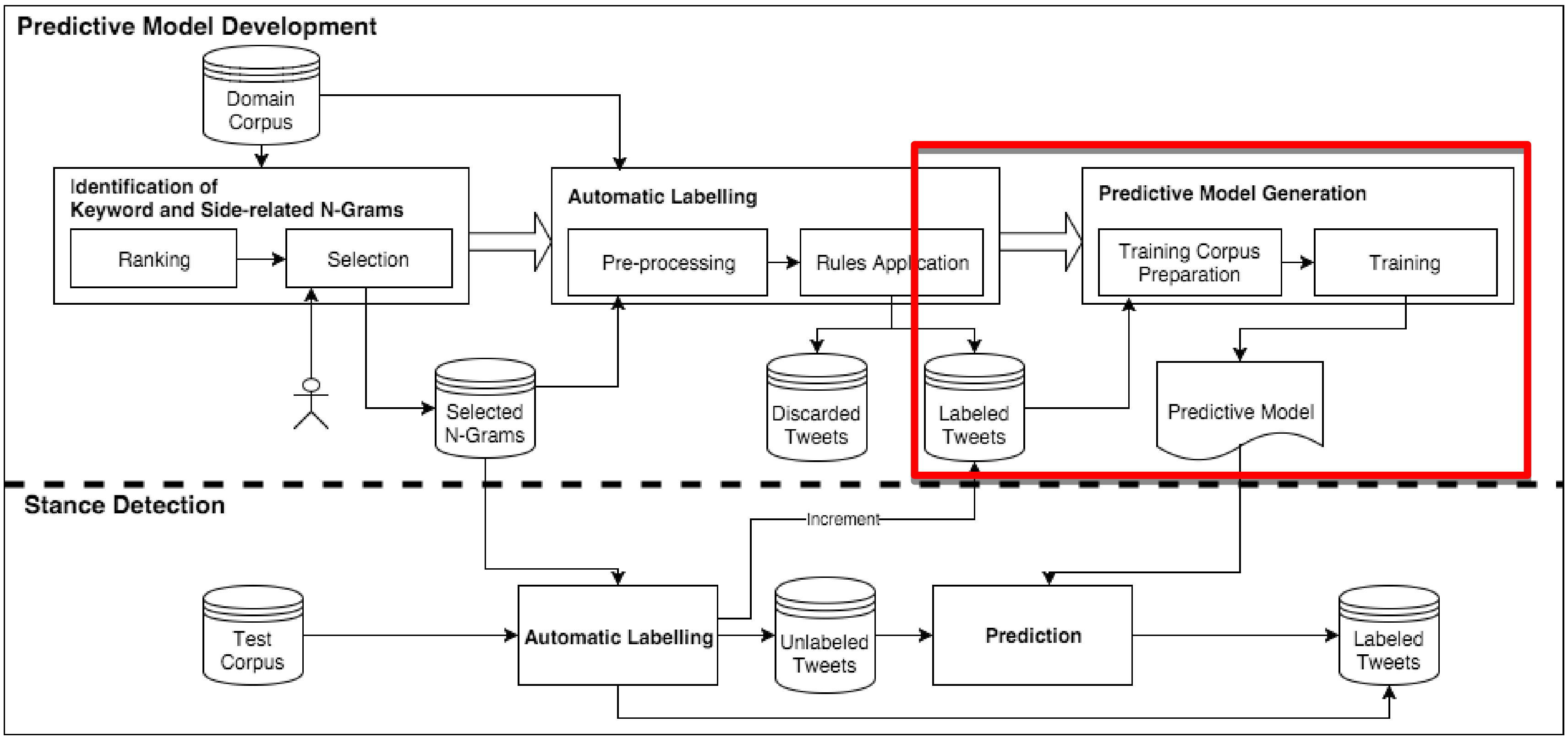
FEATURES
Presence of at least one Favor/Against Key N-grams
Presence of at least one Favor/Against Target N-grams
Presence of at least one hashtag
Tweet Polarity

Automatic Labeling

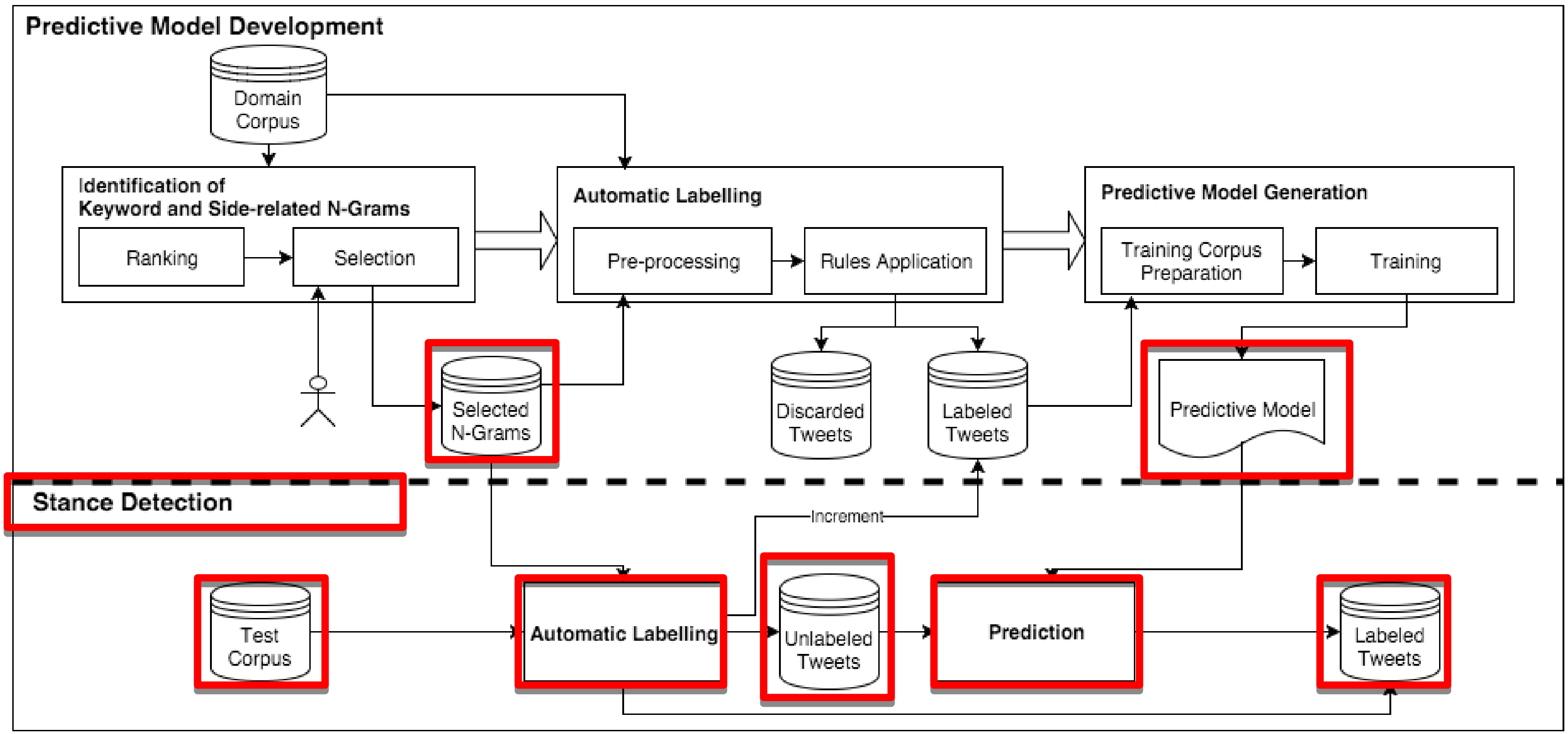


- Input: selected n-grams and a dataset
- Tweet Pre-processing
  - features extraction
  - tweet polarity detection (combination of off-the-shelf APIs)
- Rules Application
- Output: Filtered labeled tweets and discarded tweets

Predictive Model Generation



Method Overview: Stance Detection

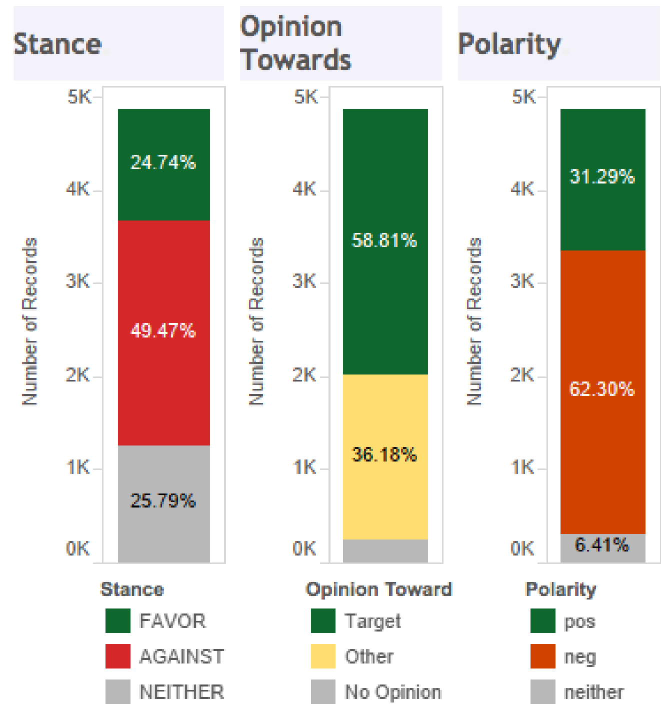


Experiments

- Goal:
  - Generality of the method for stance detection
  - 6 datasets on various domains
- Rules coverage
- Rules precision
- Stance prediction

Datasets: SemEval 2016 – Task 6

- Stance: Against, Favor or None
- Subtask A – Supervised
  - 5 targets with 2 datasets each (training and test)
  - Atheism, Climate change is a real concern, Feminism, Hillary Clinton and Legalization of Abortion
- Subtask B – Semi-supervised/Unsupervised
  - 1 targets with 2 datasets each (domain and test)
  - Donald Trump



Fonte: http://www.saifmohammad.com/WebPages/StanceDataset.htm

.inf INSTITUTO DE INFORMÁTICA UFRGS

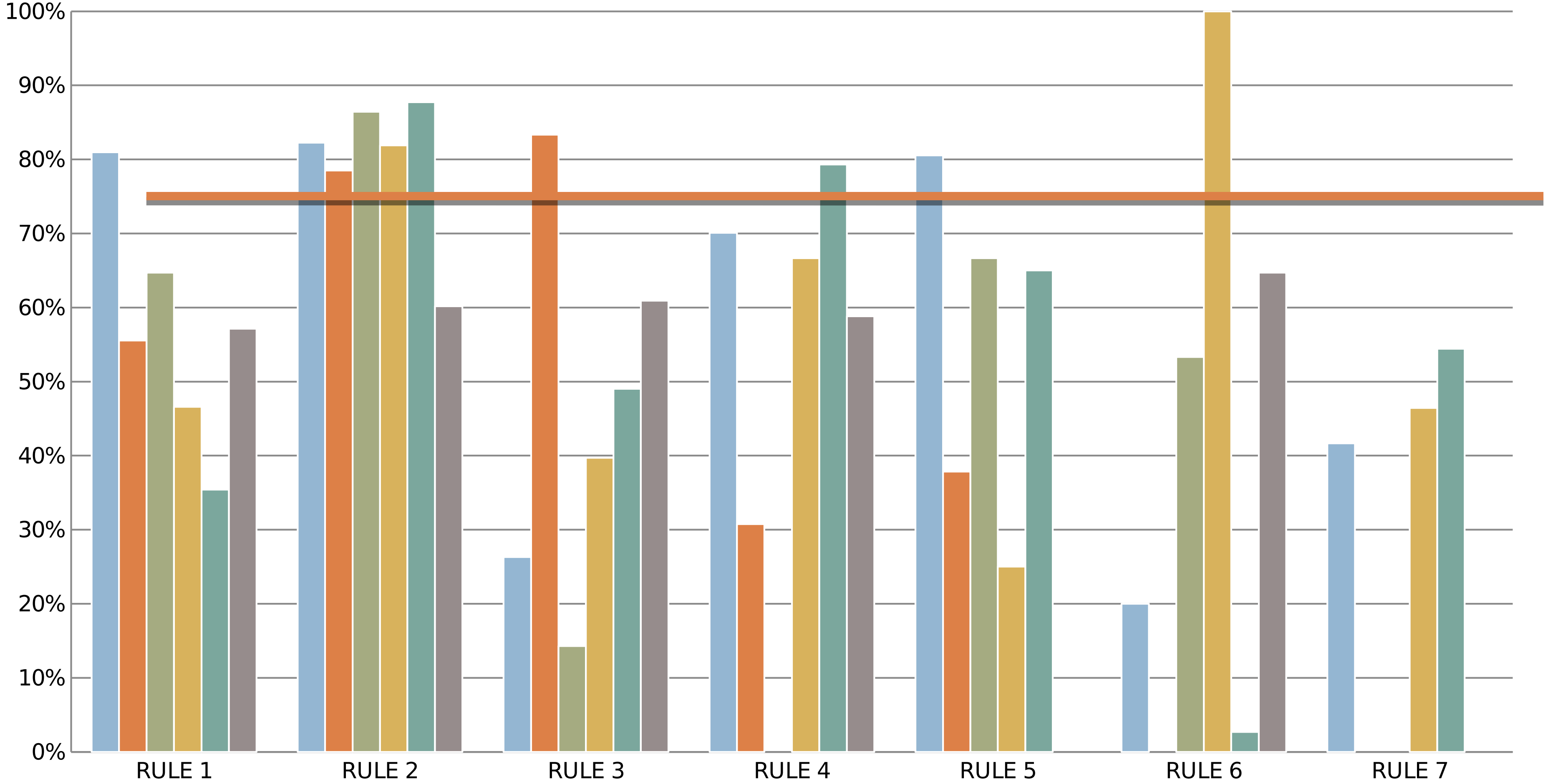
Rules Coverage

- Average corpus coverage: 75%
- In general, Rules 2, 3, 4 and 7 were representative
  - 13% to 17%
- Rules 5 and 6 are representative only for Atheism
- Rule 1 is representative only for Feminism

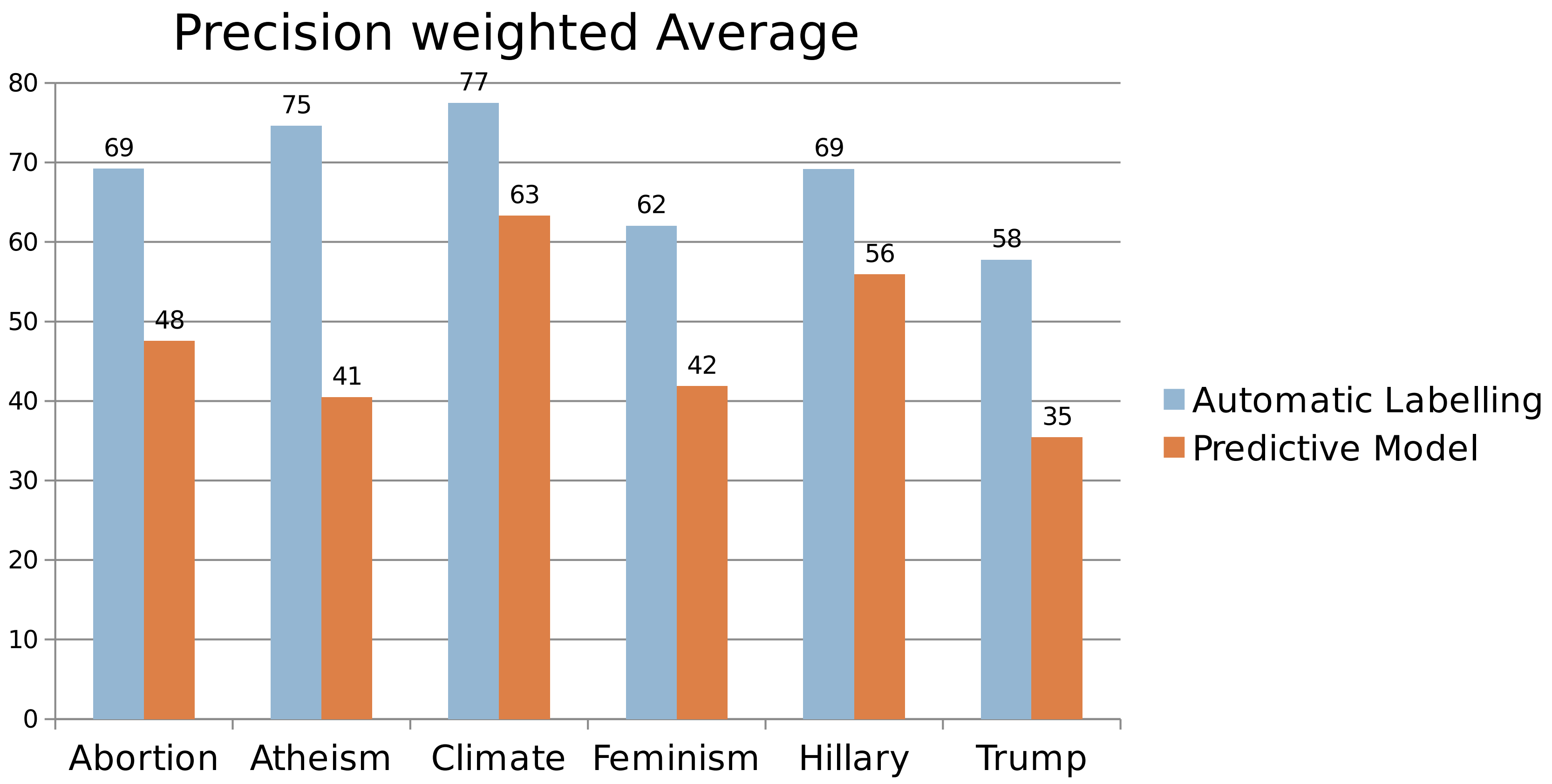
Rule	Abortion		Atheism		Climate		Feminism		Hillary		Trump		% Average
	Nb.	%	Nb.	%	Nb.	%	Nb.	%	Nb.	%	Nb.	%	
1-Keyword-Favor	21	2	36	5	17	3	103	11	48	5	1	0	4.38
2-Keyword-Against	248	28	107	15		0	105	11	163	17	1392	7	12.99
3-Favor-Positive	19	2		0	118	21	136	14	155	17	6120	30	13.93
4-Favor-Negative	97	11		0	42	7	288	30	174	19	7551	37	17.32
5-Against-Positive	36	4	186	25	5	1	4	0	20	2	18	0	5.50
6-Against-Negative	55	6	156	21	18	3	3	0	37	4	44	0	5.83
7-None	168	19	111	15	120	21	183	19	158	17	171	1	15.41
Discarded	239	27	137	19	244	43	129	14	179	19	5372	26	24.63
Total	883		733		564		949		934		20669		

INSTITUTO
DE INFORMÁTICA
UFRGS

Rules Precision

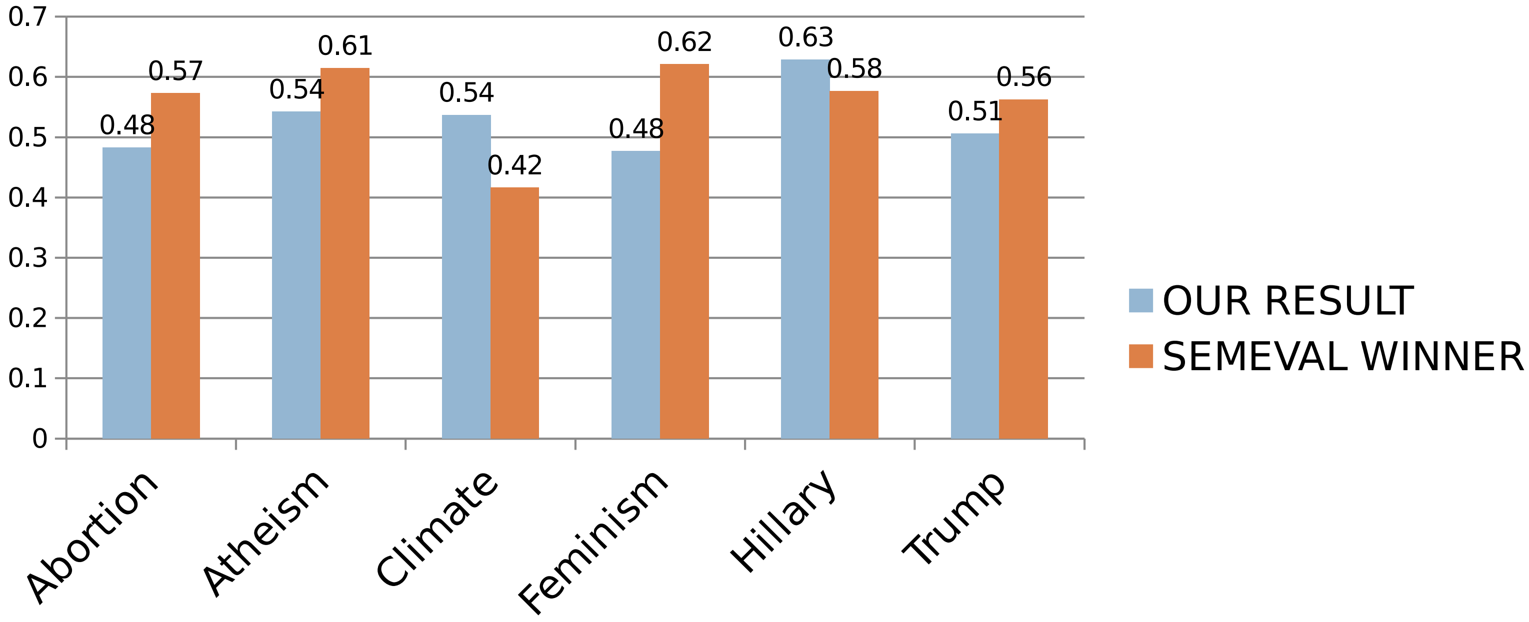


Automatic Labeling x Predictive Model



Results x Baseline

$$F_{average} = \frac{F_{favor} + F_{against}}{2}$$



Except for Trump, all the baselines were developed using a supervised metho

Strengths and Weakness

- Strengths
  - Simplicity of the method
  - May be applied to different domains/targets
  - Simplify the manual corpus annotation effort
    - Restricted to n-grams
- Weakness
  - Dependent on the appropriate selection of n-grams
    - Requires domain knowledge
  - Some rules do not perform well
  - Performance depends on the prevalence of the class

Future Work

- Key and target N-grams automatic identification
- Revised set of rules
- Neutral stance identification improvement
- Improvement of supervised-learning predictive models
  - Predictive model features
  - Automatic extraction of training instances from authority twitter profiles
  - Classification algorithms or committees