Automatic Classification and Entity Relation Detection in Hungarian Spinal MRI Reports

András Kicsi, Klaudia Szabó Ledenyi, Péter Pusztai and László Vidács

University of Szeged, Department of Sofware Engineering

MTA-SZTE Research Group on Artificial Intelligence

Motivation

- A great number of radiologic reports, usually free-text

  - These contain extractable knowledge

- A number of potential uses: quality assurance, quick comprehension, comparison to previous conditions, statistics, training data for deep learning applications

Overview
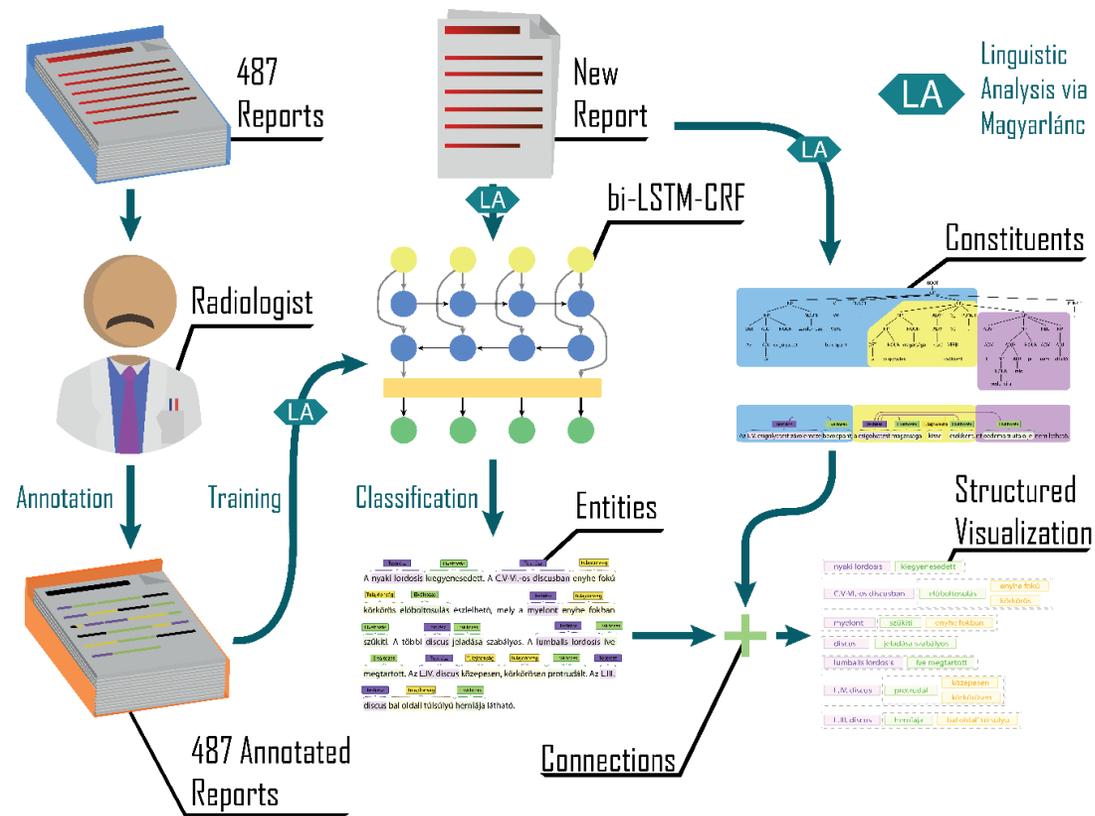
Annotations

♦ 487 Hungarian spinal reports were annotated manually

♦ Annotated entities:
anatomic locations, disorders, properties

The [Location]L2-L3 [Location]disk is preserved. Enhancing [Property]peridural [Location]fibrosis [Disorder] noted at [Location]L2-L3 level [Property]mildly deforming [Disorder] the [Location]thecal sac with [Property]dominant [Property]extrinsic [Disorder]impression on the [Property]right lateral [Location]thecal [Location]sac. Non [Property]enhancing [Disorder]cystic foci noted along the [Property]posterior elements representing [Property]small [Disorder]pseudomeningoceles. [Property]Postoperative [Disorder]fusion and [Disorder]laminectomy noted at [Location]L4-L5 level with [Property]osseous [Disorder]fusion [Property]anteriorly. [Property]Multilevel [Location]endplate, [Location]disk and [Location]facet [Property]degenerative [Disorder]changes noted.

Classification

◆ Bi-LSTM-CRF

◆ Character embeddings, additional linguistic features

◆ The CRF layer performed the final tagging on the Bi-LSTM output

◆ Whole terms, not just tokens: inside-outside-beginning tagging
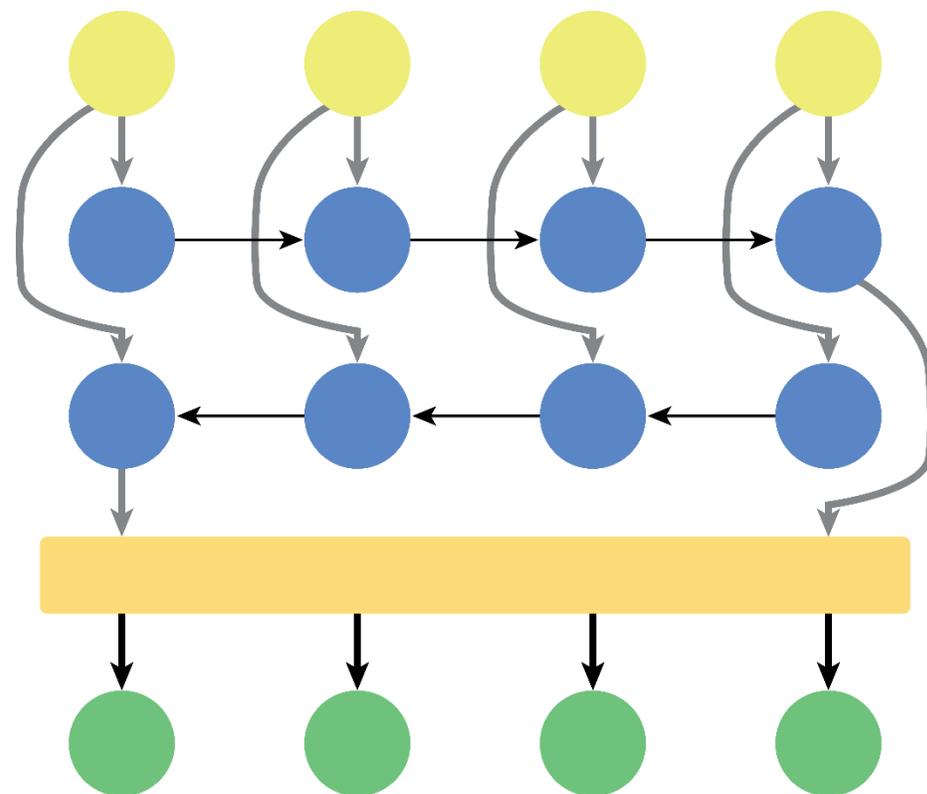
Classification Results

Class	Precision	Recall	F1-score	Support
B-Disorder	0.9105	0.9111	0.9108	1 608
I-Disorder	0.8475	0.8589	0.8532	893
B-Location	0.9518	0.9329	0.9422	1 311
I-Location	0.9404	0.9568	0.9485	1 203
B-Property	0.8712	0.8628	0.8670	729
I-Property	0.8939	0.8551	0.8741	414
Micro Avg.	0.9101	0.9076	0.9089	6 158

Connections

Az L.V. csigolyatest zárólemeze beroppant, a csigolyatest magassága kissé csökkent, itt oedemára utaló jel nem látható.

◆ The tags are great, but more semantic information is needed, entity relationships and negations are also extremely important

◆ Supported connections:
Disorder-Location, Property-Disorder, Location-Location, Disorder-Disorder

Constituents
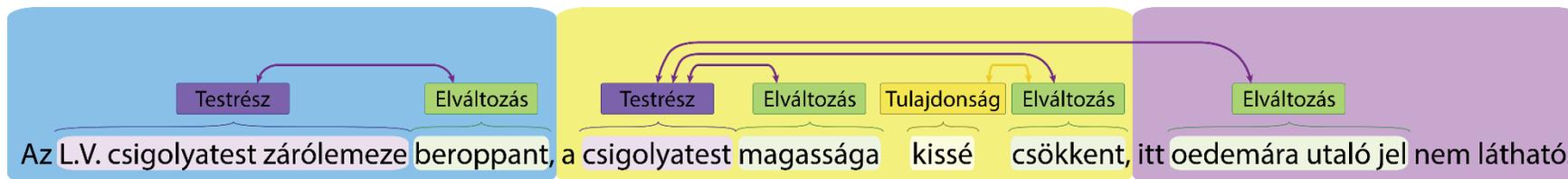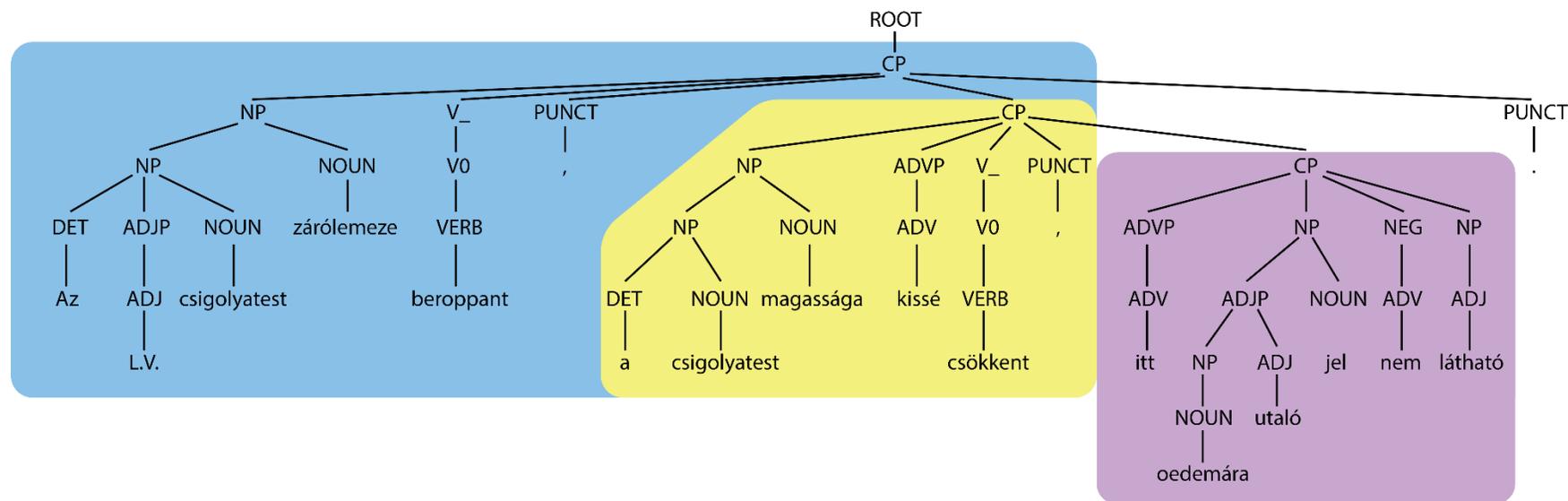
Representation

◆ The detected entities and connections are displayed in a tree structure

L2-L3 disk	preserved

peridural	fibrosis	Enhancing

L2-L3 thecal sac	deforming	mildly

thecal sac	impression	dominant
		extrinsic
		right lateral

| cystic foci | enchancing |
| | posterior |

pseudomeningoceles	small

L4-L5	fusion	Postoperative
	laminectomy	
	fusion	osseous
		anteriorly

| endplate | changes | multilevel |
| | | degenerative |

| disk | changes | multilevel |
| | | degenerative |

| facet | changes | multilevel |
| | | degenerative |

Summary

◆ The reports contain valuable information

◆ Classification based on 487 reports, locations, disorders and properties classified by Bi-LSTM-CRF

◆ Connections and negations are ascertained via linguistic information

◆ An easy-to-comprehend tree representation is constructed

◆ The extracted information will be useful in later machine learning experiments detecting disorders in images

Thank You!