

Multi-Task Learning in Deep Neural Networks at EVALITA 2018

Andrea Cimino, Lorenzo De Mattei and Felice Dell'Orletta

ItaliaNLP Lab - www.italianlp.it

Istituto di Linguistica Computazionale "Antonio Zampolli" - CNR - Pisa

Dipartimento di Informatica - Università di Pisa

EVALITA, Turin, November 12th, 2018



- ▶ In EVALITA edition the majority of task were classification (binary) tasks, among them:
 - ▶ **ABSITA** (Aspect Based Sentiment Analysis)
 - ▶ **HaSpeeDe** (Hate Speech Detection)
 - ▶ **GxG** (Gender X-Genre, author profiling in terms of gender)
 - ▶ **IronITA** (Irony and Sarcasm Detection)
- ▶ We designed and developed a general purpose system based on deep neural networks and evaluated the performance on these 4 shared tasks

- ▶ Task related issues
- ▶ Neural Network architecture adaptation for GxG
- ▶ Experiments and results performed on the **GxG 2018** shared task

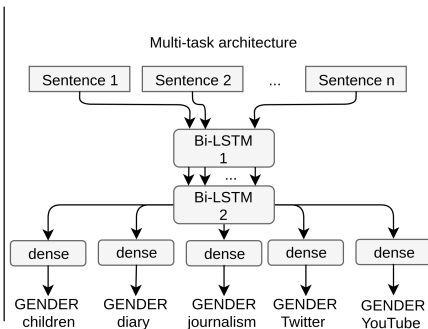
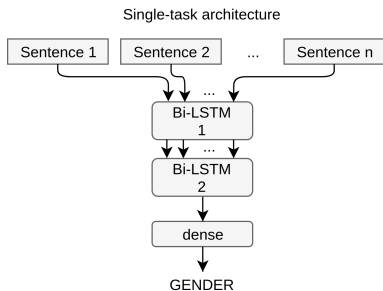
- ▶ Children Essays
- ▶ Diary
- ▶ Journalism
- ▶ Twitter
- ▶ YouTube

- ▶ Children Essays: *"...Questa estate **mi sono divertito** molto perché mio padre ha preso casa nella località del Circeo. La casa era a due piani, al piano terra c'era un giardino dove il mio gatto se la spassava. [...] Da Omar a Simone. Ciao."*
- ▶ Diary: *"**Nonno mi racconti di quando**" ... Sara, la mia nipotina la incuriosisce la mia vita, la mia infanzia così diversa e così lontana, sprofondata nel tempo e con radici così lontane da questa casa. [...] Poi attaccava all'altro consueto gancio la cesta del pane a pagnotte di granoturco, inaccessibile ai bambini."*

- ▶ Twitter: *"Secondo me i parrucchieri possiedono gli specchi che ti fanno sembrare più **bella**."*
- ▶ Twitter: *"le parole di #Rastelli "se avessimo segnato prima noi, se rosati non le avesse parate, se..se..se" mettevi a giocare c***o! #forzacasteddu"*
- ▶ YouTube: *"aspetta....sei bellissimo"*
- ▶ YouTube: *"La fanno risultare strafiga, ma è una 1/2 cessa."*

- ▶ *"Donne al lavoro, bimbi all'asilo, dobbiamo farcela DIFENDERE gli asili, ad ogni costo. "Semplicemente perché non sono un costo, ma un investimento a vantaggio di tutti, una risorsa per la collettività". Luigi Guerra non ha dubbi, [...]"*

Single-task & Multi-task architectures



Classification results of the different learning models on the official test set in terms of accuracy for the **in-domain** tasks (* marks runs that outperformed all the systems that participated to the task).

Model	CH	DI	JO	TW	YT	Avg
SVMa	0.545	0.514	0.475	0.539	0.585	0.532
SVM	0.550	0.649	0.555	0.567	0.555*	0.575
STL	0.545	0.541	0.500	0.595*	0.512	0.539
MTL	0.640*	0.676*	0.470	0.561	0.546	0.579
Avg	0.57	0.595	0.500	0.538	0.550	0.551

Classification results of the different learning models on the official test set in terms of accuracy for the **cross-domain** tasks. (* marks runs that outperformed all the systems that participated to the task).

Model	CH	DI	JO	TW	YT	Avg
SVM	0.540	0.514	0.505	0.586	0.513*	0.532
STL	0.640*	0.554	0.495	0.609*	0.510	0.562
MTL	0.535	0.595	0.510	0.500	0.500	0.528
Avg	0.572	0.554	0.503	0.565	0.508	0.541

Thanks for your attention!
Questions?