What is GMU and How to use it in my Thesis

a novel module that combines multiple sources of information, which is optimized with respect to the end goal objective function.

# Gates

module is based on the idea of gates for selecting which parts of the input are more likely to contribute for correctly generating the desired output.

input-dependent gate-activation pattern that determines how each modality contribute to the output of hidden units.

# GATED MULTIMODAL UNIT FOR MULTIMODAL FUSION

Multimodal learning is closely related to data fusion. Data fusion looks for optimal ways of combining different information sources into an integrated representation that provides more information than the individual sources (Bhatt & Kankanhalli, 2011). This fusion can be performed at different levels, that can be categorized into two broad categories:

1. feature fusion (early fusion)

Feature fusion, also called early fusion, looks for a subset of features from different modalities, or combinations of them, that better represent the information needed to solve a particular problem.

2. decision fusion

On the other hand, decision fusion, or late fusion, combines decisions from different systems, e.g. classi- fiers, to produce consensus. This consensus may be reached by a simple average, a voting system or a more complex Bayesian framework.

In this work we present a model, based on gated neural networks, for data fusion that combines ideas from both feature and decision fusion. The model, called Gated Multimodal Unit (GMU).

Figure 1. shows an example of a bimodal approach, where x\_v is the visual info and x\_t is the textual info. A multi-modal version can be found in the paper of GMU.

vA diagram of a flowchart

Description automatically generated

Figure 1

The equations governing this GMU are as follows:

* hv =tanh(Wv ·xv)
* ht =tanh(Wt ·xt)
* z=σ(Wz ·[xv,xt])
* h=z∗hv +(1−z)∗ht
* Θ = {Wv,Wt,Wz}

Obviously these can be integrated in ML workflows and backpropagation.