

- • • Activation fn: arctan, Norm fn: sparseMax, alpha : 0.3
- Activation fn: arctan, Norm fn: softMax, alpha : 0.8
- Activation fn: GELU, Norm fn: softMax, alpha : 0.8
- Activation fn: swish, Norm fn: softMax, alpha : 0.8
- Activation fn: arctan, Norm fn: softMax, alpha : 0.5
- • • Activation fn: GELU, Norm fn: sparseMax, alpha : 0.5
- • • Activation fn: swish, Norm fn: sparseMax, alpha : 0.5
- • • Activation fn: arctan, Norm fn: sparseMax, alpha : 0.5
- • • Activation fn: LeakyReLU, Norm fn: sparseMax, alpha : 0.5
- Activation fn: softPlus, Norm fn: softMax, alpha : 0.8
- Activation fn: ELU, Norm fn: softMax, alpha : 0.5
- • • Activation fn: softPlus, Norm fn: sparseMax, alpha : 0.3
- Activation fn: LeakyReLU, Norm fn: softMax, alpha : 0.5
- • • Activation fn: arctan, Norm fn: sparseMax, alpha : 0.8
- Activation fn: arctan, Norm fn: softMax, alpha : 0.3
- • • Activation fn: GELU, Norm fn: sparseMax, alpha : 0.3
- • • Activation fn: ELU, Norm fn: sparseMax, alpha : 0.5
- • • Activation fn: ReLU, Norm fn: sparseMax, alpha : 0.5
- • • Activation fn: arctan, Norm fn: sparseMax, alpha : 0.1
- • • Activation fn: sigmoid, Norm fn: sparseMax, alpha : 0.5
- • • Activation fn: softPlus, Norm fn: sparseMax, alpha : 0.5
- Activation fn: softPlus, Norm fn: softMax, alpha : 0.5
- Activation fn: arctan, Norm fn: softMax, alpha : 0.1
- Activation fn: ReLU, Norm fn: softMax, alpha : 0.8
- • • Activation fn: swish, Norm fn: sparseMax, alpha : 0.1
- • • Activation fn: swish, Norm fn: sparseMax, alpha : 0.8
- • • Activation fn: swish, Norm fn: sparseMax, alpha : 0.1
- • • Activation fn: ELU, Norm fn: sparseMax, alpha : 0.8
- Activation fn: ELU, Norm fn: softMax, alpha : 0.8
- Activation fn: ReLU, Norm fn: softMax, alpha : 0.5
- • • Activation fn: LeakyReLU, Norm fn: sparseMax, alpha : 0.3
- Activation fn: LeakyReLU, Norm fn: softMax, alpha : 0.8
- • • Activation fn: sigmoid, Norm fn: sparseMax, alpha : 0.3