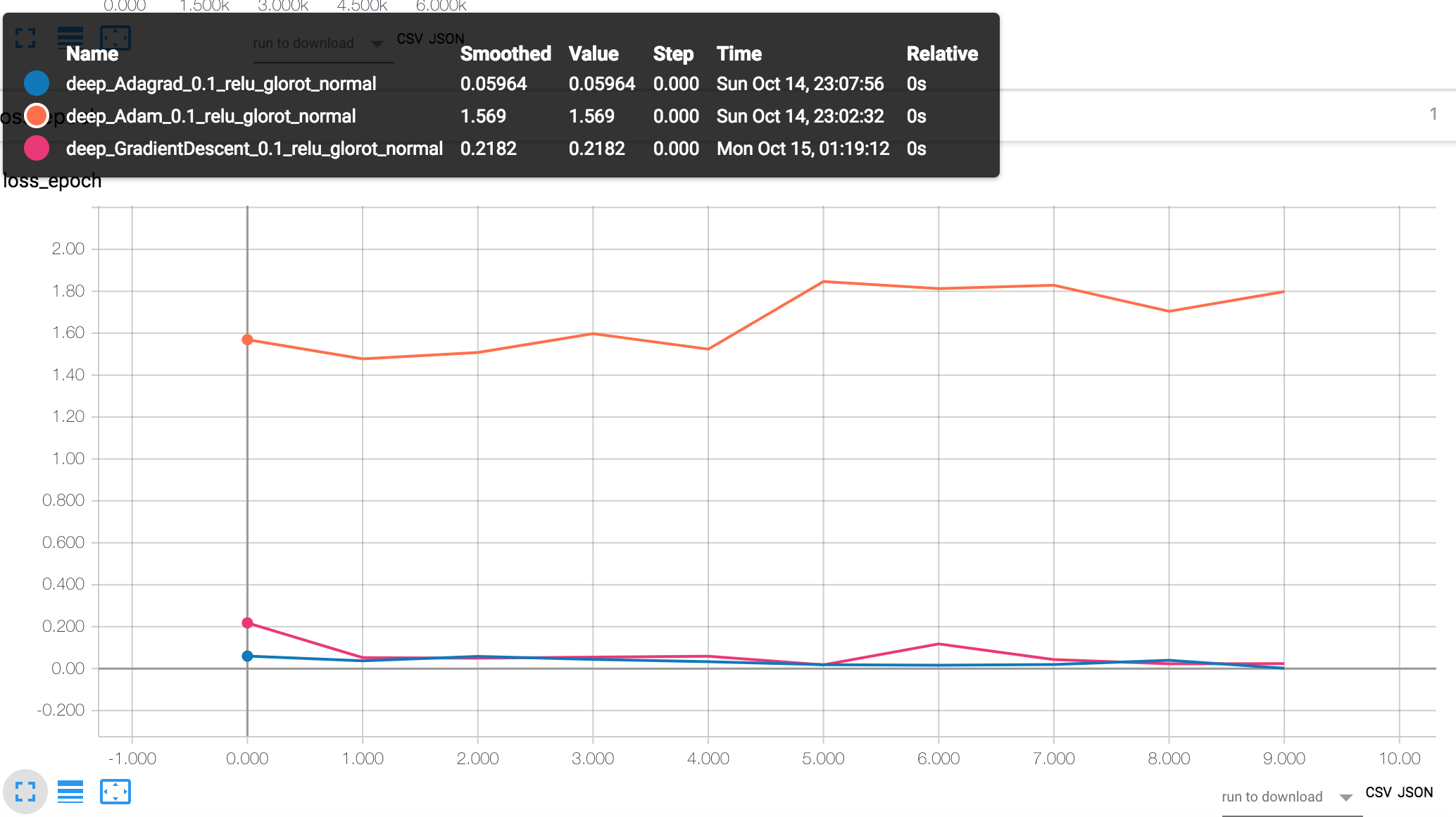
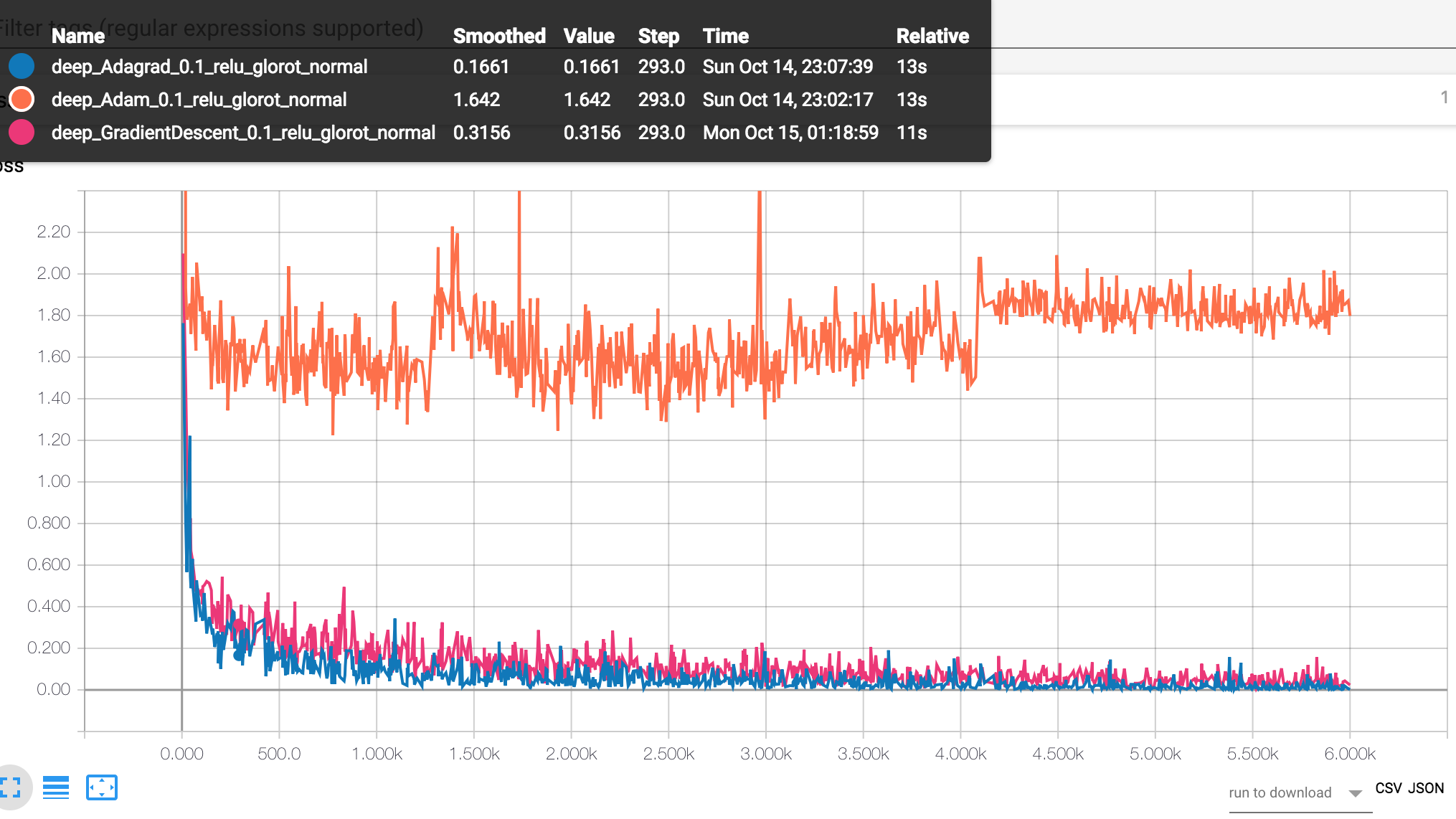
**Plotting different optimizers - using the relu activation function, Glorot Normal Initializer and learning rate of 0.1**

**Training Loss vs Epoch**

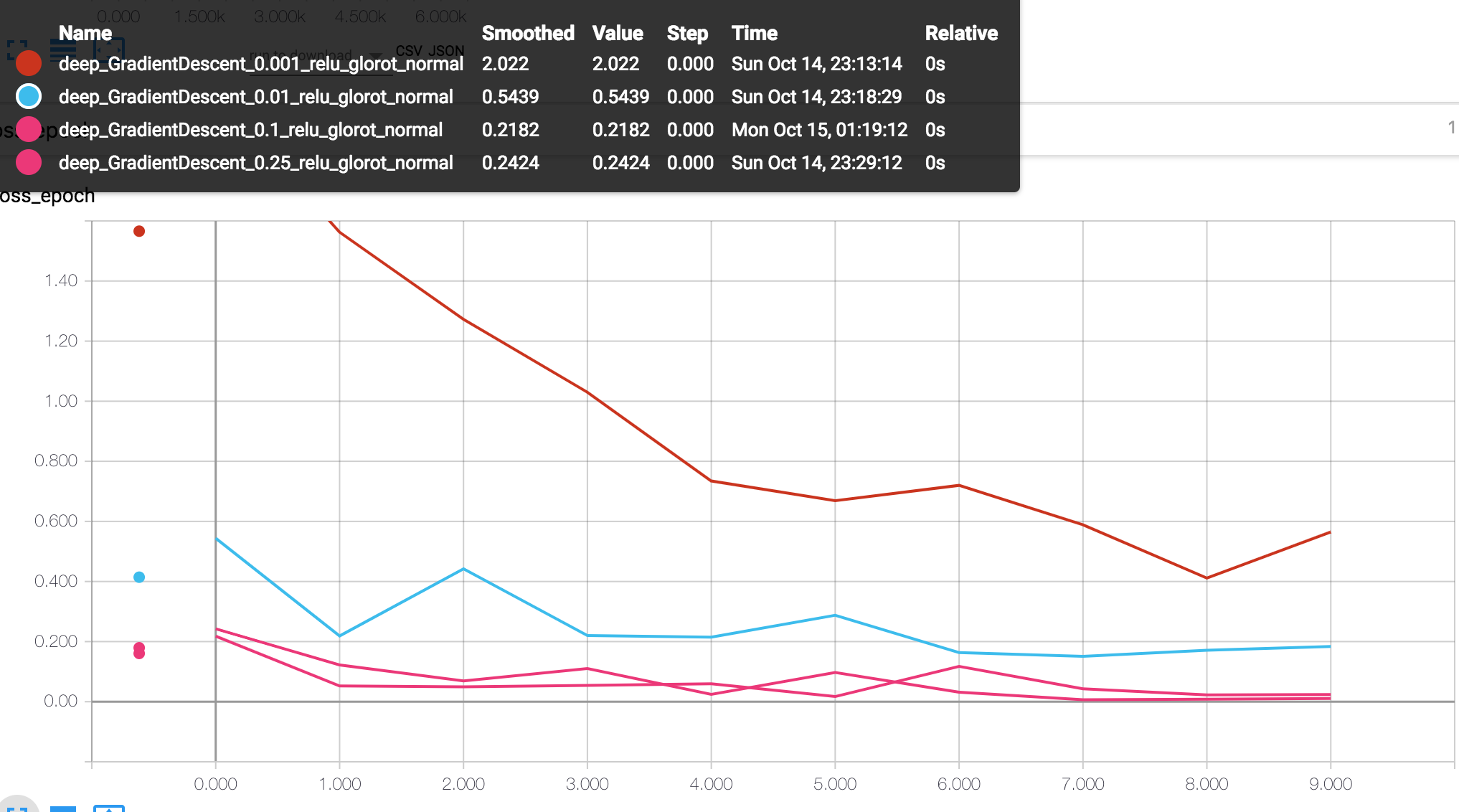


**Training Loss vs Step**

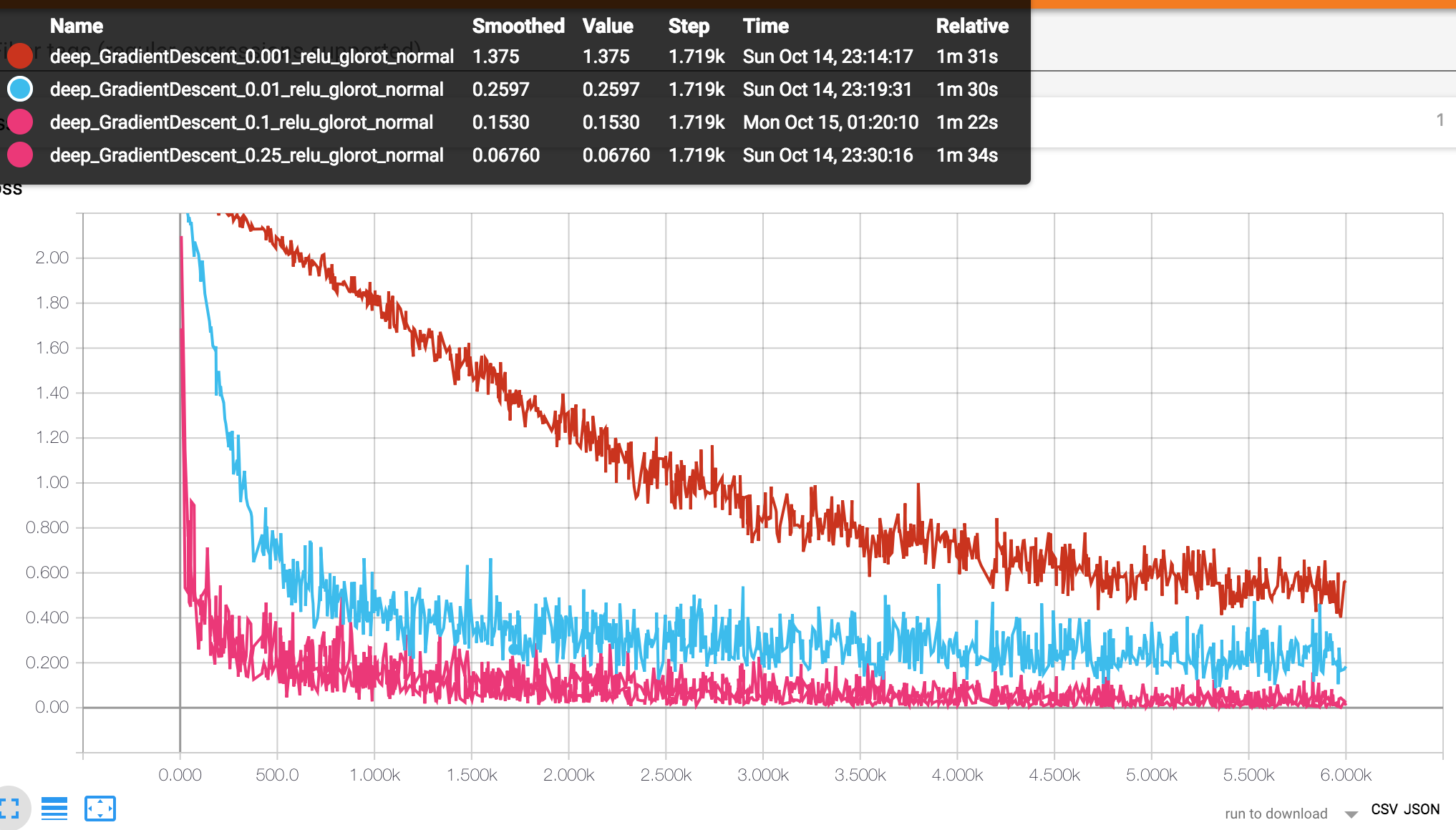


**Plotting different learning rates - using the relu activation function, Gradient Descent Optimizer and Glorot Normal Initializer**

**Training Loss vs Epoch**

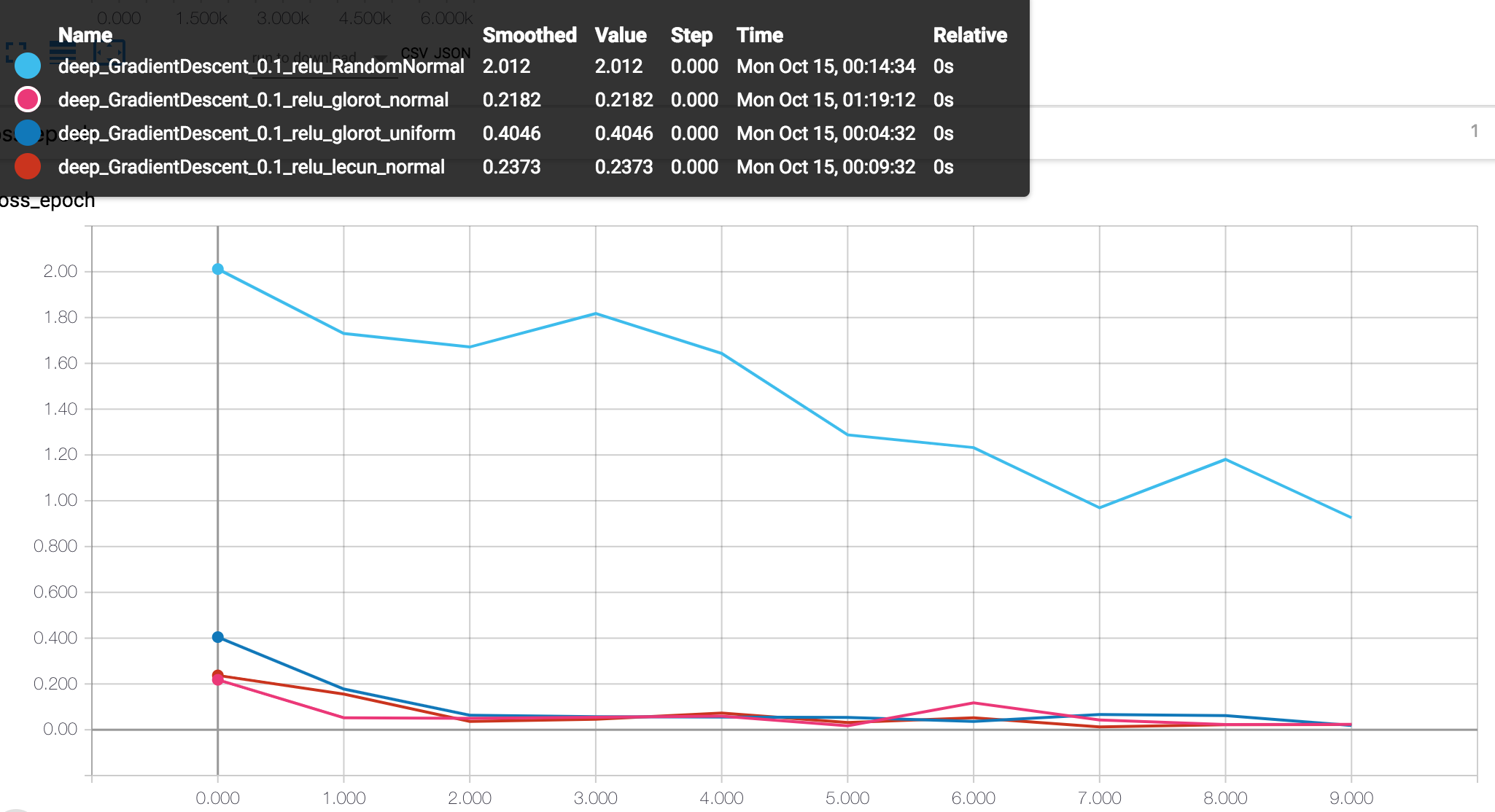


**Training Loss vs Step**

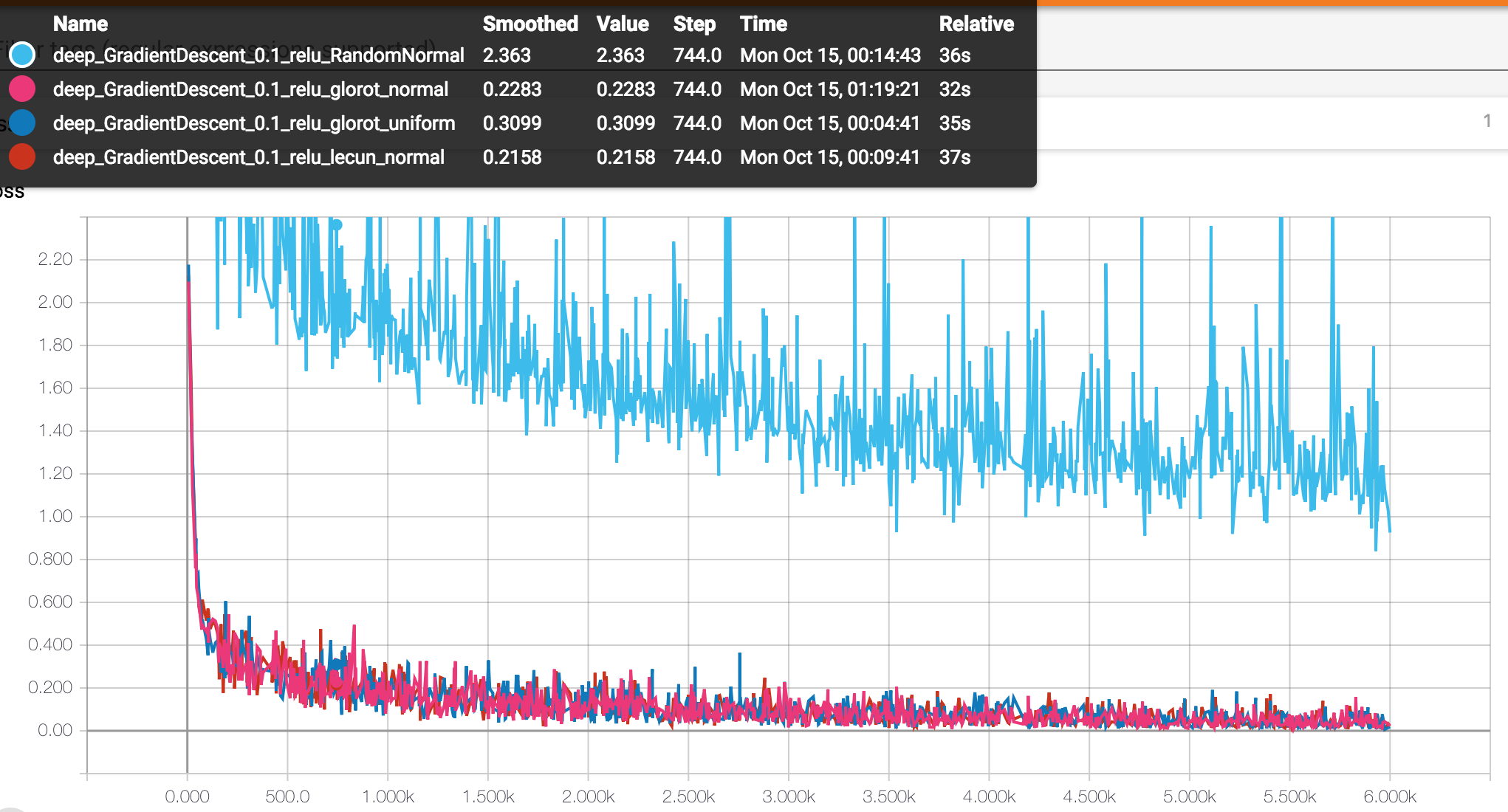


**Plotting different initializers - using the relu activation function, Gradient Descent Optimizer** **and learning rate of 0.1**

**Training Loss vs Epoch**

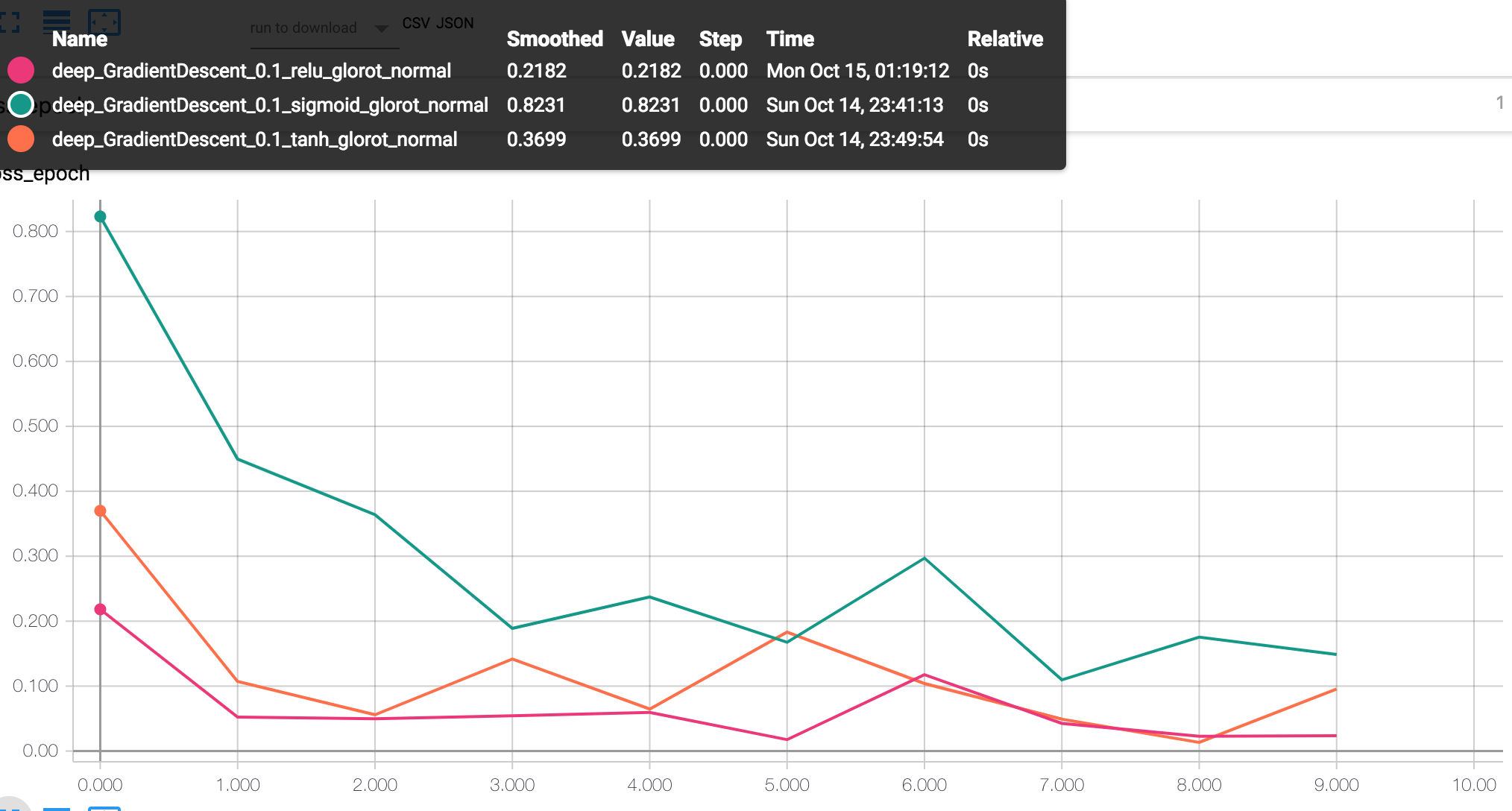


**Training Loss vs Step**

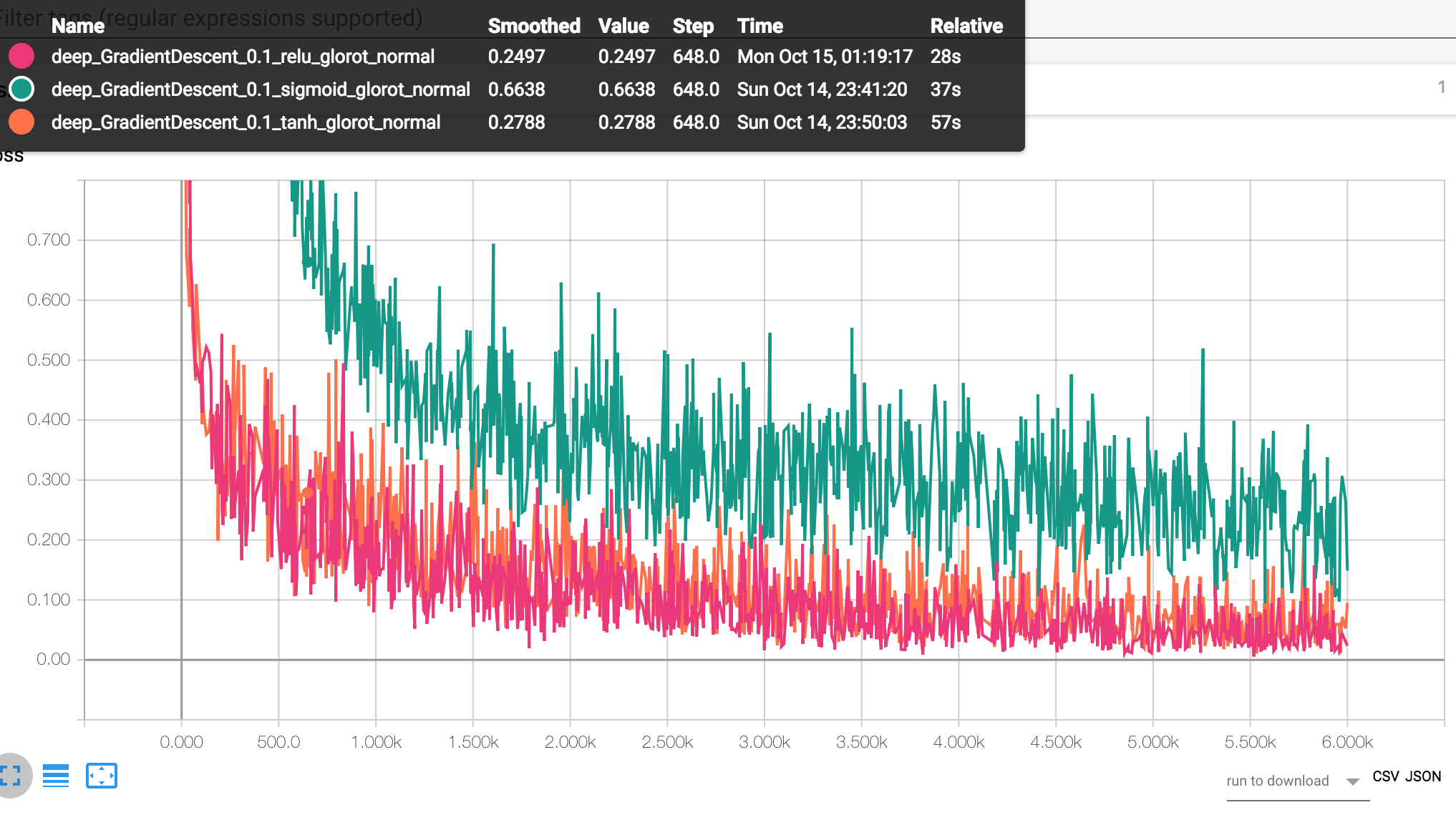


**Plotting different activation functions - using a learning rate of 0.1, Gradient Descent Optimizer and Glorot Normal Initializer**

**Training Loss vs Epoch**

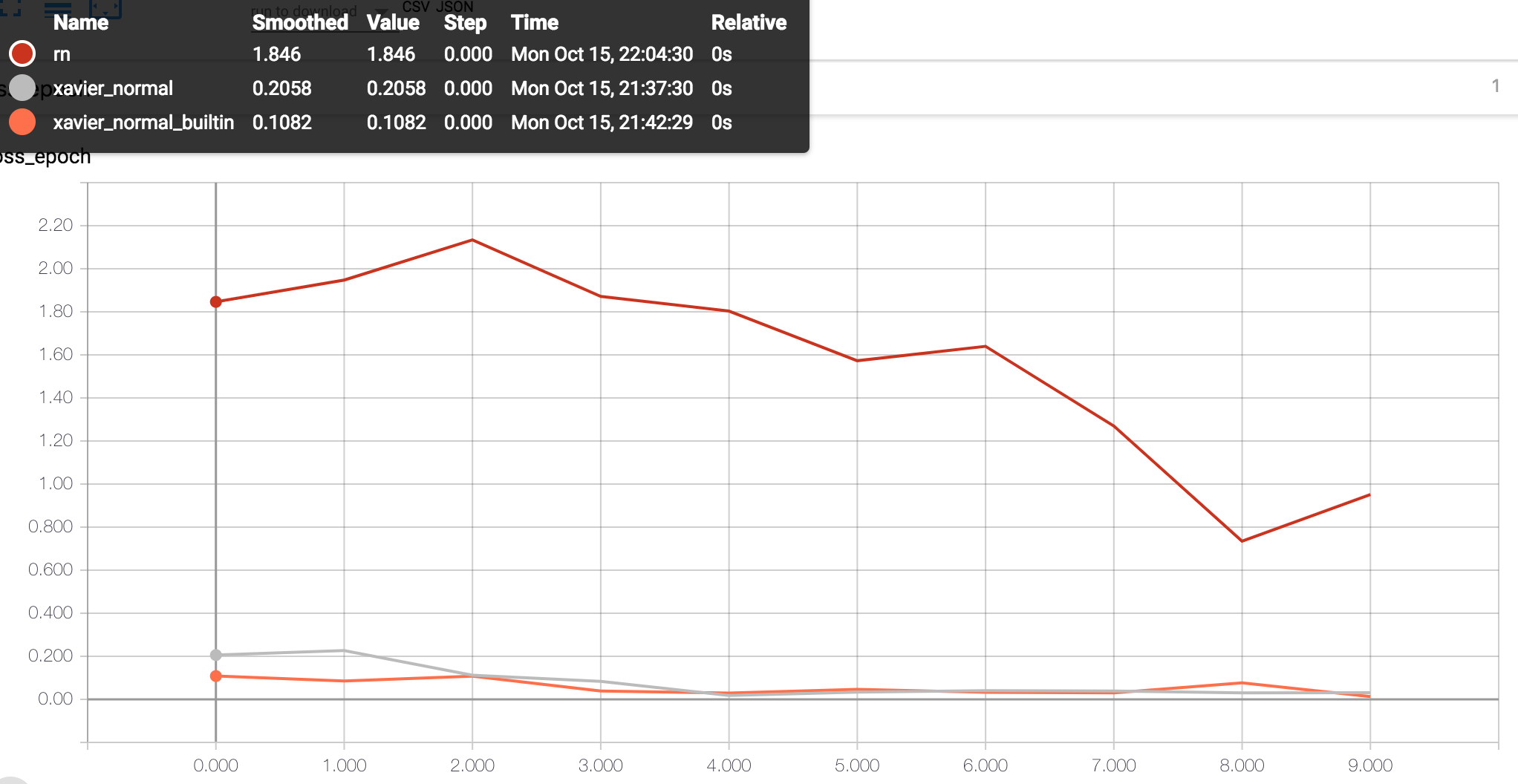


**Training Loss vs Step**

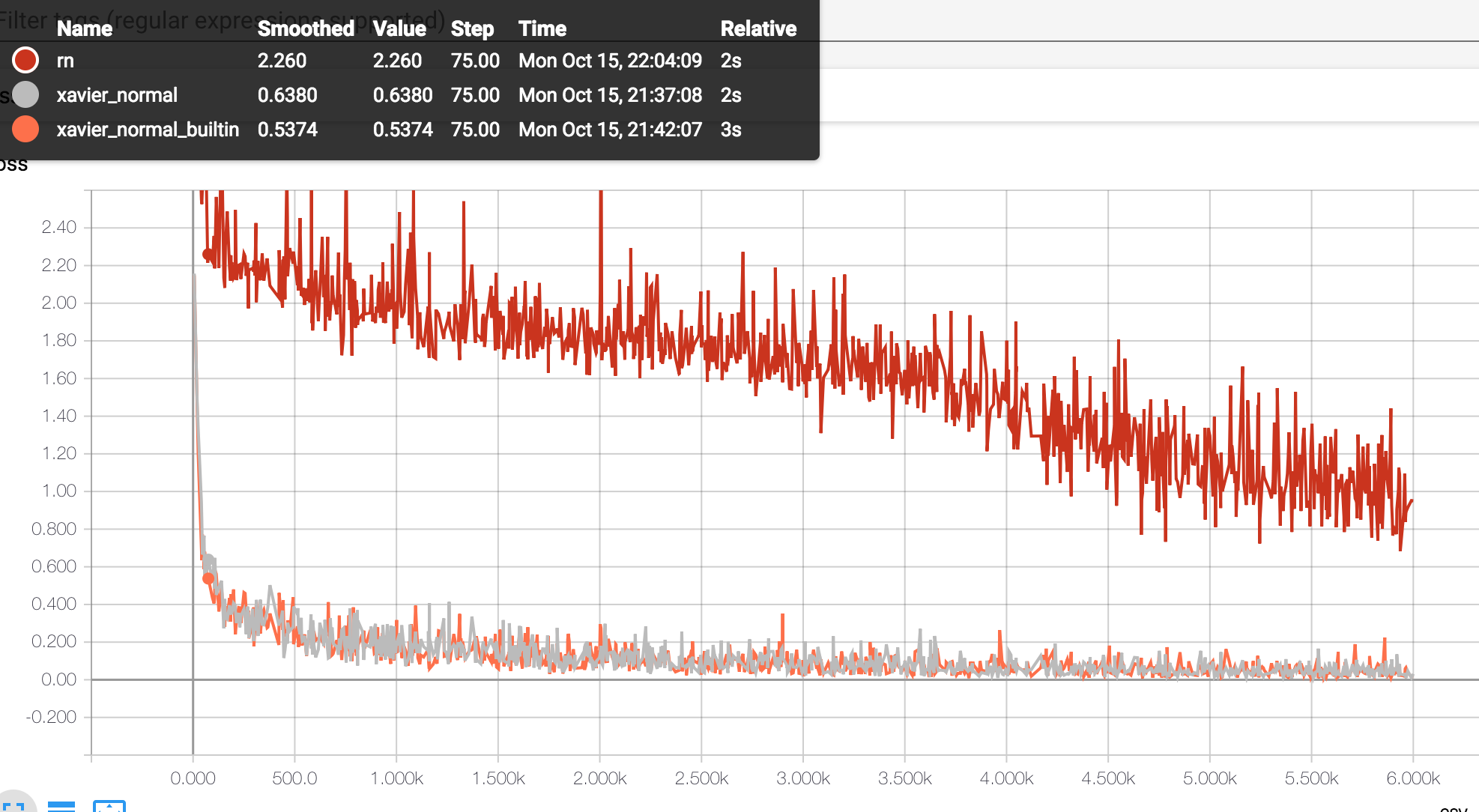


**Comparison of the Custom Xavier Initialization with Built-in Xavier Initialization and Random Normal Initialization**

**Training Loss vs Epoch**

****

**Training Loss vs Step**

****