Uuuuufff, that thing is steep!!!

This connects to the same topic as on Friday. Background: I work at Selecta AG and want to make a ML-model to predict how many products are needed to be refilled on the next day. In most cases, you only have to refill one or two products of a sort per day, or even none. The problem is, that these 3 categories apply to so many cases, that the model doesn't try to predict refill amounts bigger than 2.

My questions:

- The model is designed to be a regression. Could a classification model perform better? If so, what optimizer / loss function are the best?
- What optimizer and loss function could I try to achieve better results for the regression? (I'm using adam optimizer and mean squared error as loss function)