Relevance heatmap


$\boldsymbol{v} = [\ 0.4\ 0.5\ 0.1\ ]^{\mathsf{T}}$

A vector of the percentages of relevance scores assigned to each 28x28 block

Heatmap Quality

Explanation

$f(\boldsymbol{x})$

$$\cos(\boldsymbol{m}, \boldsymbol{v}) = \frac{\boldsymbol{m}^{T}\boldsymbol{v}}{\|\boldsymbol{m}\|\|\boldsymbol{v}\|}$$

RNN

Prediction
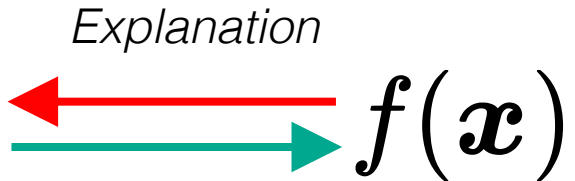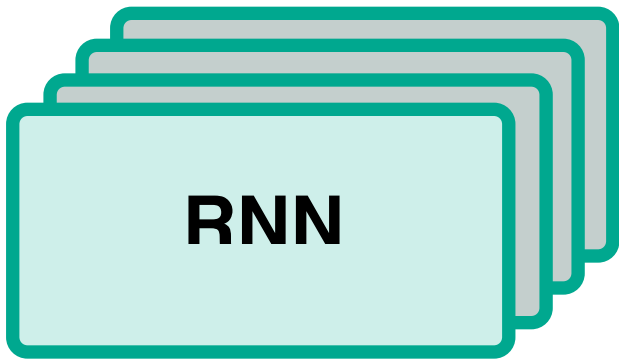
Input Sequence

$\boldsymbol{m} = [\ 1\ 1\ 0\ ]^{\mathsf{T}}$

A vector of variables indicating whether each 28x28 block belongs to the majoirty digit