Bidirectional Timestamp-Guided Attention Model

Input MBS items
w_t t=1,...,T

Input Timestamps
τ_t t=1,...,T

MBS Embedding
1 x E
1 x V
V x E
$x_t = w_t W_e$

Bidirectional LSTM
$\overrightarrow{h}_x = LSTM(x_t)$ t=1,...,T
$\overleftarrow{h}_x = LSTM(x_t)$ t=1,...,T
$h_x = [\overrightarrow{h}_x, \overleftarrow{h}_x]$
T x H

Bidirectional LSTM
$\overrightarrow{h}_\tau = LSTM(\tau_t)$ t=1,...,T
$\overleftarrow{h}_\tau = LSTM(\tau_t)$ t=1,...,T
$h_\tau = [\overrightarrow{h}_\tau, \overleftarrow{h}_\tau]$

Dense (tanh)
$\gamma = tanh(h'_x W_x + b_x)$
H x T   T x T

Dense (tanh)
$\beta = tanh(h'_\tau W_\tau + b\tau)$

Convex combination
$\delta = \lambda \gamma + (1-\lambda) \beta$
H x T

Attention probabilities

Dense (softmax)
$\alpha = exp(\delta_t W_\alpha) / \sum_{t=1}^{T} [exp(\delta_t W_\alpha)]$
H x T   T x T

$\Omega = h_x \odot \alpha'$ - - - - → Contribution
T x H

$C = \sum_{t=1}^{T} \Omega_t x_t$ - - - - → Context
H x E

Averaging
$\tilde{c} = 1/H \sum_{j=1}^{H} C_j$
1 x E

Dense & Dropout

Output $p(y \mid w_1,...,w_T) = \sigma(W \tilde{c} + b)$