

$y \in \{\text{INSERT, NoINSERT}\}$ $y \in \{\text{PRUNE, NoPRUNE}\}$

fully-connected layer
(insert decisions)

fully-connected layer
(prune decisions)

max pool

Bi-LSTM

Learned embeddings

$\mathbf{x} = \langle s \rangle$ She said that $\langle c \rangle$ the protocol limits $\langle \text{PRUNEdobj?} \rangle$ fire risk $\langle / \text{PRUNEdobj?} \rangle \langle / c \rangle \langle / s \rangle$

