Figure 1: The full pipeline of our model (for $k = 0.5$), leveraging several stacks of interleaved convolutional/pooling layers (that, unlike DiffPool, *drop* rather than aggregate nodes), as well as a JK-net-style summary, combining information at different scales.