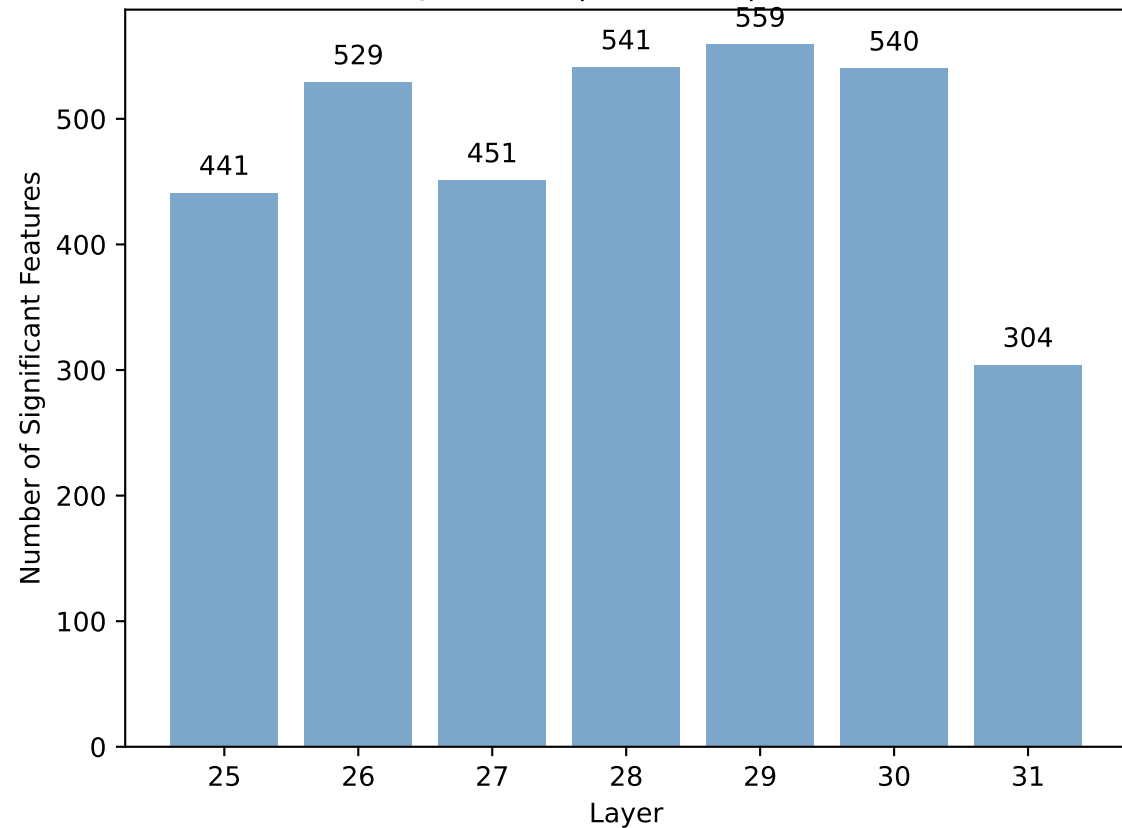**Real Feature Distribution Across Layers**
(p<0.001, |Cohen's d|>0.3)

Feature Direction by Layer
(Real Data from 3,365 Features)

Legend:
- Safe-promoting (Cohen's d > 0)
- Risk-promoting (Cohen's d < 0)