$$Z_0 = [X_class; X_p^1\mathbf{E}; X_p^2\mathbf{E}; ...X_p^N\mathbf{E}] + \mathbf{E}_pos, \mathbf{E} \in \mathbb{R}^{(p^2,c)D}, \mathbf{E} \in \mathbb{R}^{(N+1)D}$$
$$Z_\ell^{'} = MSA(LN(Z_{\ell-1})) + Z_{\ell-1}\ell = 1...L$$
$$Z_\ell = MLP(LN(Z_\ell^{'})) + Z_\ell^{'}\ell = 1...L \tag{1}$$
$$Y = LN(Z_L^0)$$