In-context learning	Input prompt	Desired Output
Natural language processing	berry, baya, apple, manzana, banana	plátano
	Japan, mochi, France, croissant, Greece	baklava
Supervised learning $y_i = f(x_i) + \text{noise}$	$x_1, y_1, x_2, \ldots, x_{i-1}, y_{i-1}, x_i$	$f(x_i)$
Dynamical systems $x_{i+1} = f(x_i) + \text{noise}$	$x_1, x_2, x_3, \ldots, x_{i-2}, x_{i-1}, x_i$	$f(x_i)$

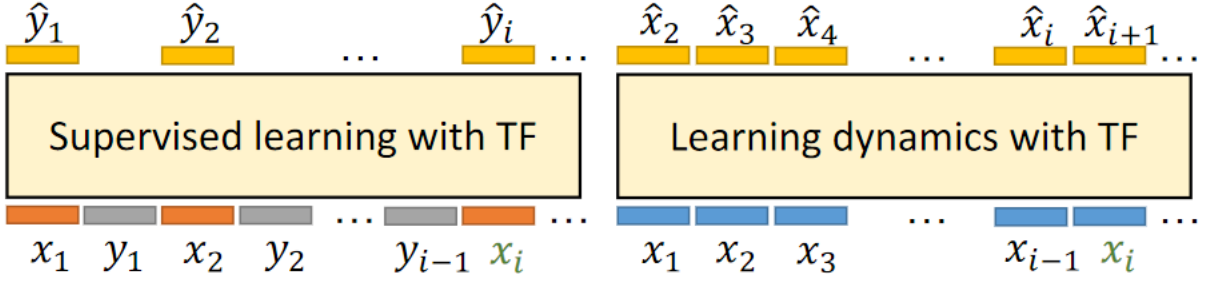

Figure 1: Examples of in-context learning. We focus on the lower two settings in the table where a transformer admits a supervised dataset or dynamical system trajectory as a prompt. Then, it auto-regressively predicts the output following an input example $\boldsymbol{x}_i$ based on the prompt $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i)$.