

Column Projection in “WHERE” Clause

Zihua

Task Description

url	Category	food	temperature	resting_time
https://www.foodsafety.gov/keep/charts/mintemp.html	Ground Meat & Meat Mixtures	Beef, Pork, Veal, Lamb	160	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Ground Meat & Meat Mixtures	Turkey, Chicken	165	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Fresh Beef, Veal, Lamb	Steaks, roasts, chops	145	3 minutes
https://www.foodsafety.gov/keep/charts/mintemp.html	Poultry	Chicken & Turkey, whole	165	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Poultry	Poultry breasts, roasts	165	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Poultry	Poultry thighs, legs, wings	165	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Poultry	Duck & Goose	165	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Poultry	Stuffing (cooked alone or in bird)	165	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Pork and Ham	Fresh pork	145	3 minutes
https://www.foodsafety.gov/keep/charts/mintemp.html	Pork and Ham	Fresh ham (raw)	145	3 minutes
https://www.foodsafety.gov/keep/charts/mintemp.html	Pork and Ham	Precooked ham (to reheat)	140	None
https://www.foodsafety.gov/keep/charts/mintemp.html	Eggs & Egg Dishes		Cook until yolk	None

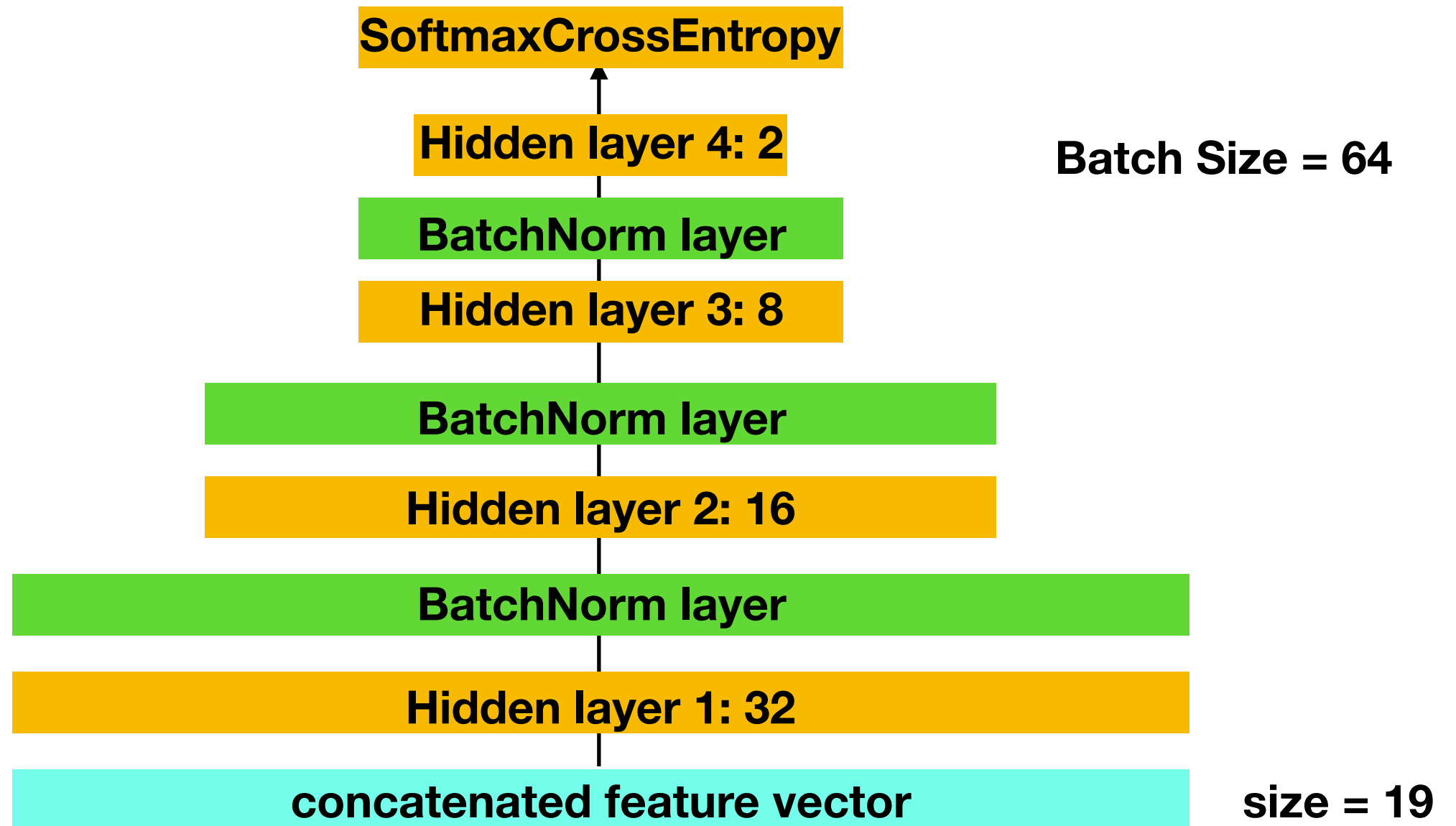
SELECT “temperature” FROM food_safety WHERE “food” ~ “chicken”

Features

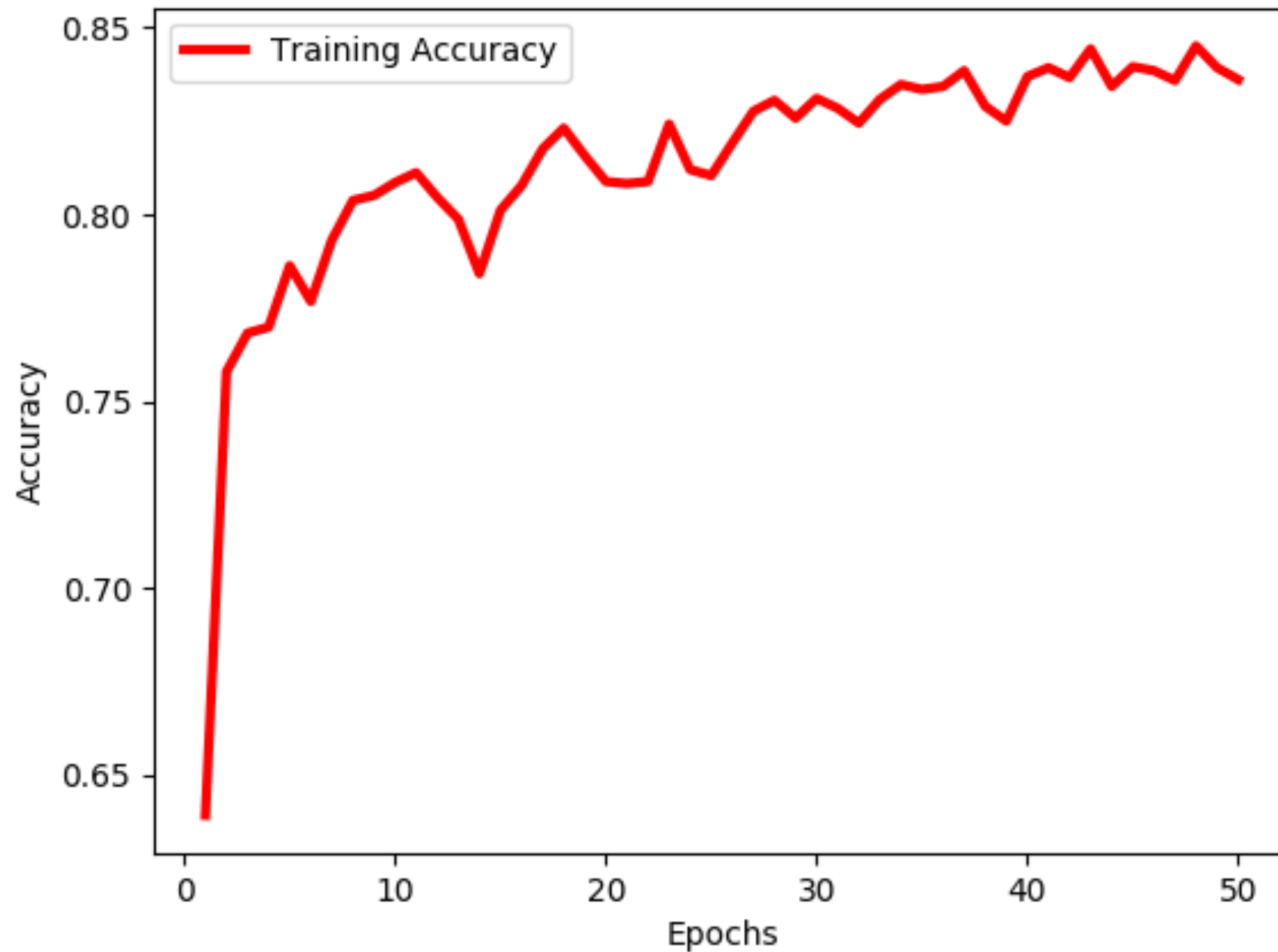
- Number of columns (Lucile)
- Word2vec proximities (Lucile)
- Column Types (Zihua)
- Question Types (Nithin) —> Transform to one-hot vector

ML Model

For training set, there are total 2046 cases, 115 positive (label = 1), 1931 negative (label = 0), so copy positive cases 16 times



Training Accuracy



Column projection

Training set

	Postitive	Negative
True	105	1457
False	10	474
	115	1931

Test set

	Postitive	Negative
True	25	422
False	5	107
	30	529