

Table S1. Baseline data of HNSCC patients in the training, validation, and testing cohorts.

Variables		Training set (%)	Validation set (%)	Testing set (%)
Total		457	104	182
Age (mean $\pm$ SD)		58.3 $\pm$ 12.2	59.9 $\pm$ 12.7	53.2 $\pm$ 10.4
Sex	Male	271 (59.3)	54 (51.9)	166 (91.2)
	Female	186 (40.7)	50 (48.1)	16 (8.8)
Site	Tongue	249 (54.5)	49 (47.1)	113 (62.1)
	Gingiva	72 (15.8)	13 (12.5)	14 (7.7)
	Buccal	49 (10.7)	14 (13.5)	41 (22.5)
	Floor of mouth	49 (10.7)	10 (9.6)	5 (2.7)
	Lip	14 (3.1)	10 (9.6)	5 (2.7)
	Palate	14 (3.1)	4 (3.8)	1 (0.5)
	Jaw	7 (1.5)	4 (3.8)	1 (0.5)
	Oropharynx	3 (0.7)	0 (0.0)	2 (1.1)
Grade	Well differentiated	340 (74.4)	86 (82.7)	165 (90.7)
	Moderately differentiated	116 (25.4)	17 (16.3)	16 (8.8)
	Poorly differentiated	1 (0.2)	1 (1.0)	1 (0.5)
PD-L1 expression	CPS < 1	95 (20.8)	17 (16.3)	29 (15.9)
	CPS $\geq$ 1	362 (79.2)	87 (83.7)	153 (84.1)

Table S2. Performances of machine learning models in the validation and testing datasets.

Model name	Dataset	AUC (95% CI)	Sensitivity	Specificity	Accuracy	PPV	NPV
LR	Validation	0.900 (0.832 - 0.967)	0.685	>0.999	0.731	>0.999	0.349
	Testing	0.851 (0.789 - 0.913)	0.609	>0.999	0.665	>0.999	0.299
NaiveBayes	Validation	0.705 (0.569 - 0.841)	0.955	>0.999	0.885	0.914	0.636
	Testing	0.697 (0.588 - 0.807)	0.833	0.933	0.791	0.915	0.350
SVM	Validation	0.888 (0.797 - 0.979)	0.843	0.867	0.846	0.974	0.481
	Testing	0.825 (0.752 - 0.898)	0.705	0.846	0.725	0.965	0.324
RandomForest	Validation	0.775 (0.643 - 0.907)	0.910	>0.999	0.875	0.942	0.556
	Testing	0.738 (0.648 - 0.827)	0.538	>0.999	0.593	0.977	0.250
ExtraTrees	Validation	0.843 (0.723 - 0.964)	0.809	>0.999	0.808	0.960	0.414
	Testing	0.824 (0.747 - 0.900)	0.660	0.880	0.687	0.963	0.293
XGBoost	Validation	0.842 (0.724 - 0.960)	0.888	0.923	0.875	0.963	0.545
	Testing	0.796 (0.719 - 0.873)	0.583	0.923	0.632	0.978	0.270
LightGBM	Validation	0.860 (0.734 - 0.986)	0.933	0.800	0.913	0.965	0.667
	Testing	0.736 (0.611 - 0.860)	0.776	0.731	0.769	0.945	0.352
AdaBoost	Validation	0.845 (0.735 - 0.954)	0.865	0.800	0.856	0.962	0.500
	Testing	0.815 (0.739 - 0.890)	0.647	0.923	0.687	0.981	0.304