

Figure S1 Calibration curves of the XGB, DT, RF and SVM models in the (A) training set and (B) validation set. XGB, eXtreme gradient boost; DT, decision tree; RF, random forest; SVM, support vector machine.

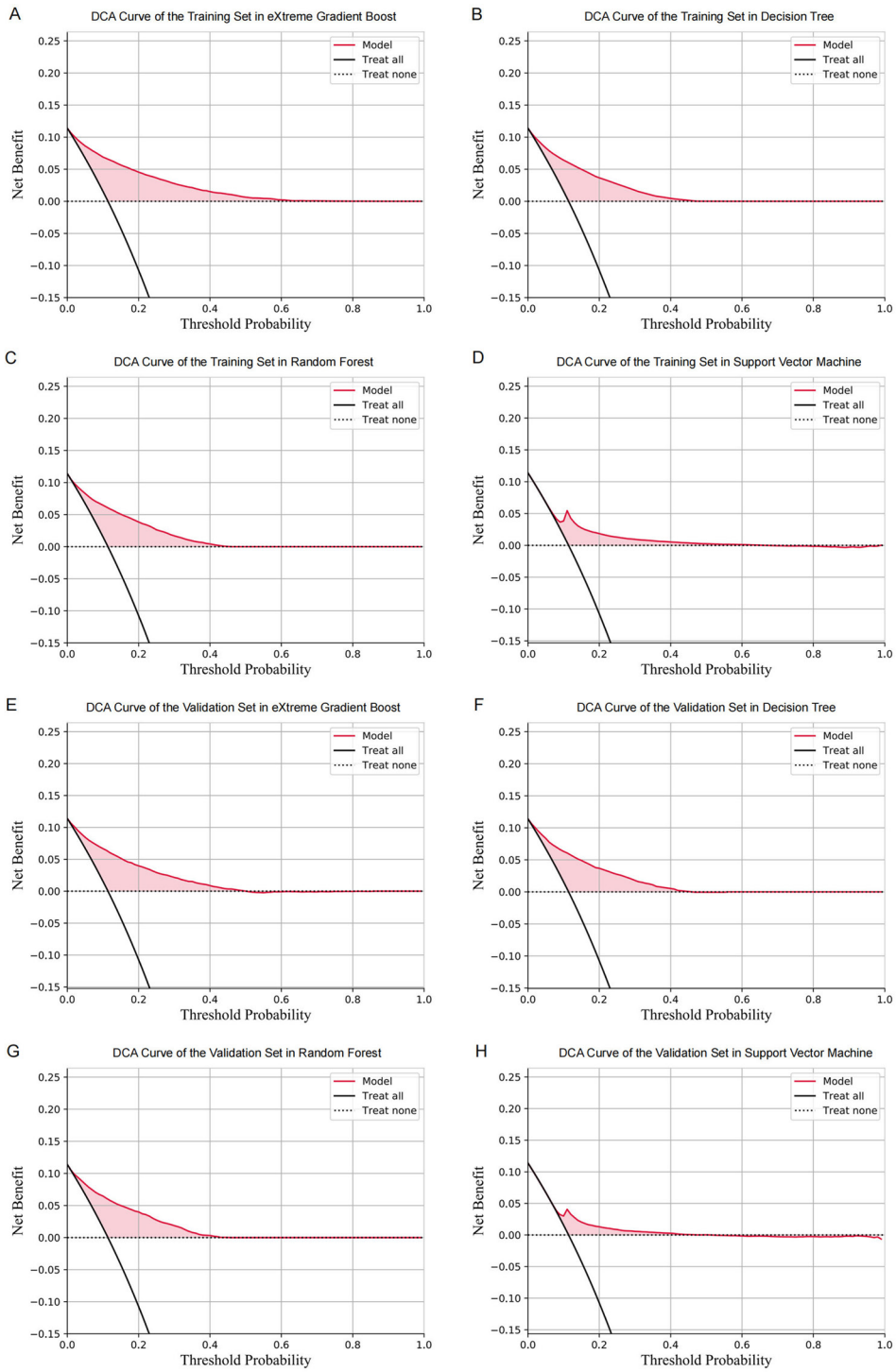


Figure S2 DCA curves of the (A) XGB model, (B) DT model, (C) RF model and (D) SVM model in the training set. DCA curves of the (E) XGB model, (F) DT model, (G) RF model and (H) SVM model in the validation set. DCA, decision curve analysis; XGB, eXtreme gradient boost; DT, decision tree; RF, random forest; SVM, support vector machine.

Table S1 Clinical characteristics of colorectal cancer patients in the training set

Characteristic	Liver metastasis (–), N=32,022	Liver metastasis (+), N=4,120	Total (N=36,142)	P value
Age at diagnosis, n (%), years				<0.001
0–20	18 (0.06)	1 (0.02)	19 (0.05)	
21–40	1,277 (3.99)	231 (5.61)	1,508 (4.17)	
41–60	10,812 (33.76)	1,730 (41.99)	12,542 (34.70)	
61–80	14,899 (46.53)	1,783 (43.28)	16,682 (46.16)	
>80	5,016 (15.66)	375 (9.10)	5,391 (14.92)	
Race, n (%)				<0.001
Black	3,678 (11.49)	653 (15.85)	4,331 (11.98)	
White	24,984 (78.02)	3,088 (74.95)	28,072 (77.67)	
Other	3,360 (10.49)	379 (9.20)	3,739 (10.35)	
Sex, n (%)				<0.001
Female	15,658 (48.90)	1,826 (44.32)	17,484 (48.38)	
Male	16,364 (51.10)	2,294 (55.68)	18,658 (51.62)	
Primary site, n (%)				<0.001
Ascending colon	5,272 (16.46)	591 (14.34)	5,863 (16.22)	
Cecum	5,897 (18.42)	842 (20.44)	6,739 (18.65)	
Descending colon	1,539 (4.81)	232 (5.63)	1,771 (4.90)	
Hepatic flexure	1,242 (3.88)	132 (3.20)	1,374 (3.80)	
Rectosigmoid junction	2,621 (8.18)	386 (9.37)	3,007 (8.32)	
Rectum	5,636 (17.60)	418 (10.15)	6,054 (16.75)	
Sigmoid colon	6,577 (20.54)	1,085 (26.33)	7,662 (21.20)	
Splenic flexure	842 (2.63)	138 (3.35)	980 (2.71)	
Transverse colon	2,396 (7.48)	296 (7.18)	2,692 (7.45)	
Grade, n (%)				<0.001
1	2,358 (7.36)	148 (3.59)	2,506 (6.93)	
2	23,669 (73.91)	2,879 (69.88)	26,548 (73.45)	
3	4,958 (15.48)	887 (21.53)	5,845 (16.17)	
4	1,037 (3.24)	206 (5.00)	1,243 (3.44)	
Histologic, n (%)				<0.001
Adenocarcinoma	23,652 (73.86)	3,390 (82.28)	27,042 (74.82)	
Adenocarcinoma in adenomatous polyp	2,040 (6.37)	140 (3.40)	2,180 (6.03)	
Squamous cell carcinoma	14 (0.04)	4 (0.10)	18 (0.05)	
Other	6,316 (19.72)	586 (14.22)	6,902 (19.10)	
T stage, n (%)				<0.001
1	3,048 (9.52)	52 (1.26)	3,100 (8.58)	
2	5,083 (15.87)	116 (2.82)	5,199 (14.38)	
3	18,716 (58.45)	2,393 (58.08)	21,109 (58.41)	
4	5,175 (16.16)	1,559 (37.84)	6,734 (18.63)	
N stage, n (%)				<0.001
0	17,812 (55.62)	680 (16.50)	18,492 (51.16)	
1	9,269 (28.95)	1,575 (38.23)	10,844 (30.00)	
2	4,941 (15.43)	1,865 (45.27)	6,806 (18.83)	
Tumor size, n (%), cm				<0.001
0.1–2.0	4,082 (12.75)	142 (3.45)	4,224 (11.69)	
2.1–4.0	10,947 (34.19)	1,120 (27.18)	12,067 (33.39)	
4.1–6.0	9,627 (30.06)	1,646 (39.95)	11,273 (31.19)	
6.1–8.0	4,497 (14.04)	766 (18.59)	5,263 (14.56)	
8.1–10.0	1,864 (5.82)	288 (6.99)	2,152 (5.95)	
>10.0	1,005 (3.14)	158 (3.83)	1,163 (3.22)	
CEA, n (%)				<0.001
Negative	20,036 (62.57)	764 (18.54)	20,800 (57.55)	
Positive	11,986 (37.43)	3,356 (81.46)	15,342 (42.45)	
Tumor deposits, n (%)				<0.001
Negative	28,052 (87.60)	2,563 (62.21)	30,615 (84.71)	
Positive	3,970 (12.40)	1,557 (37.79)	5,527 (15.29)	
Perineural invasion, n (%)				<0.001
Negative	28,322 (88.45)	2,785 (67.60)	31,107 (86.07)	
Positive	3,700 (11.55)	1,335 (32.40)	5,035 (13.93)	

CEA, carcinoembryonic antigen.

Table S2 Clinical characteristics of colorectal cancer patients in the validation set

Characteristic	Liver metastasis (-), N=13,724	Liver metastasis (+), N=1,766	Total (N=15,490)	P value
Age at diagnosis, n (%), years				<0.001
0–20	2 (0.01)	0 (0.00)	2 (0.01)	
21–40	539 (3.93)	101 (5.72)	640 (4.13)	
41–60	4,644 (33.84)	741 (41.96)	5,385 (34.76)	
61–80	6,382 (46.50)	774 (43.83)	7,156 (46.20)	
>80	2,157 (15.72)	150 (8.49)	2,307 (14.89)	
Race, n (%)				<0.001
Black	1,562 (11.38)	261 (14.78)	1,823 (11.77)	
White	10,671 (77.75)	1,338 (75.76)	12,009 (77.53)	
Other	1,491 (10.86)	167 (9.46)	1,658 (10.70)	
Sex, n (%)				0.031
Female	6,719 (48.96)	816 (46.21)	7,535 (48.64)	
Male	7,005 (51.04)	950 (53.79)	7,955 (51.36)	
Primary site, n (%)				<0.001
Ascending colon	2,297 (16.74)	242 (13.70)	2,539 (16.39)	
Cecum	2,581 (18.81)	369 (20.89)	2,950 (19.04)	
Descending colon	635 (4.63)	95 (5.38)	730 (4.71)	
Hepatic flexure	509 (3.71)	66 (3.74)	575 (3.71)	
Rectosigmoid junction	1,143 (8.33)	151 (8.55)	1,294 (8.35)	
Rectum	2,392 (17.43)	172 (9.74)	2,564 (16.55)	
Sigmoid colon	2,751 (20.05)	490 (27.75)	3,241 (20.92)	
Splenic flexure	337 (2.46)	50 (2.83)	387 (2.50)	
Transverse colon	1,079 (7.86)	131 (7.42)	1,210 (7.81)	
Grade, n (%)				<0.001
1	1,024 (7.46)	63 (3.57)	1,087 (7.02)	
2	10,099 (73.59)	1,231 (69.71)	11,330 (73.14)	
3	2,154 (15.70)	379 (21.46)	2,533 (16.35)	
4	447 (3.26)	93 (5.27)	540 (3.49)	
Histologic, n (%)				<0.001
Adenocarcinoma	10,128 (73.80)	1,451 (82.16)	11,579 (74.75)	
Adenocarcinoma in adenomatous polyp	818 (5.96)	53 (3.00)	871 (5.62)	
Squamous cell carcinoma	4 (0.03)	2 (0.11)	6 (0.04)	
Other	2,774 (20.21)	260 (14.72)	3,034 (19.59)	
T stage, n (%)				<0.001
1	1,314 (9.57)	21 (1.19)	1,335 (8.62)	
2	2,286 (16.66)	40 (2.27)	2,326 (15.02)	
3	7,933 (57.80)	1,060 (60.02)	8,993 (58.06)	
4	2,191 (15.96)	645 (36.52)	2,836 (18.31)	
N stage, n (%)				<0.001
0	7,745 (56.43)	277 (15.69)	8,022 (51.79)	
1	3,912 (28.50)	685 (38.79)	4,597 (29.68)	
2	2,067 (15.06)	804 (45.53)	2,871 (18.53)	
Tumor size, n (%), cm				<0.001
0.1–2.0	1,777 (12.95)	46 (2.60)	1,823 (11.77)	
2.1–4.0	4,702 (34.26)	478 (27.07)	5,180 (33.44)	
4.1–6.0	4,121 (30.03)	678 (38.39)	4,799 (30.98)	
6.1–8.0	1,907 (13.90)	340 (19.25)	2,247 (14.51)	
8.1–10.0	793 (5.78)	156 (8.83)	949 (6.13)	
>10.0	424 (3.09)	68 (3.85)	492 (3.18)	
CEA, n (%)				<0.001
Negative	8,448 (61.56)	310 (17.55)	8,758 (56.54)	
Positive	5,276 (38.44)	1,456 (82.45)	6,732 (43.46)	
Tumor deposits, n (%)				<0.001
Negative	12,027 (87.63)	1,109 (62.80)	13,136 (84.80)	
Positive	1,697 (12.37)	657 (37.20)	2,354 (15.20)	
Perineural invasion, n (%)				<0.001
Negative	12,110 (88.24)	1,186 (67.16)	13,296 (85.84)	
Positive	1,614 (11.76)	580 (32.84)	2,194 (14.16)	

CEA, carcinoembryonic antigen.