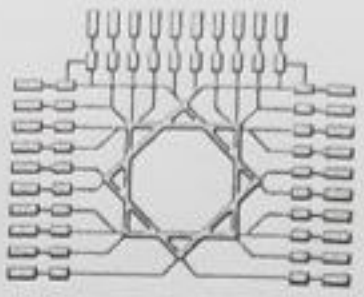


المعهد العالي للعلوم التطبيقية والتكنولوجيا
امتحان مقرر مقدمة إلى الذكاء الصنعي للعام الدراسي 2022-2023
الفصل الأول / السنة الرابعة معلوماتية
المدة: ساعتان / المادة: مفتوحة / يسمح باستعمال الآلات الحاسبة

الجزء الأول (ساعة واحدة – 50 علامة)

السؤال الأول (20 علامة): أجب عن كلّ من الأسئلة التالية بإيجاز

1) ما الفرق بين الموسطات parameters والموسطات الفوقية Hyper parameters؟ أعط مثالاً لكلٍ منهما في حالة خوارزمية تعلم أشجار القرار ID3؟
2) برأيك هل هناك تكافؤ بين استراتيجية التدريب اعتماداً على تقسيم المعطيات إلى قسمين متساويين (50% تدريب، و 50% اختبار)، واستراتيجية التدريب وفق مبدأ K-Fold Cross-validation مع قيمة K=2. برر إجابتك موضحاً نقاط التشابه والاختلاف لكل حالة؟
3) من أجل مسألة التصنيف الثنائي، ما الفرق بين نوعي الأخطاء FP و FN؟ متى يكون تقليل FP أكثر أهمية في التطبيقات العملية؟ أعط مثالاً على ذلك.. أعد المطلوب من أجل FN.
4) ما الفرق بين الطرائق المشرف عليها والطرائق غير المشرف عليها لانتقاء السمات. أيهما يؤدي إلى سمات أكثر جودة؟ برر إجابتك.

السؤال الثاني (30 علامة): نهدف في هذه المسألة إلى استخدام خوارزمية التعلم ID3 لتوقع خطر الحوادث اعتماداً على السمات التالية: time عدد السنوات منذ الحصول على إجازة القيادة، gender جنس السائق، area مكان القيادة. جرى جمع المعطيات المبينة في الجدول المرفق لإنجاز عملية التدريب. والمطلوب:

ID	time	gender	area	risk
1	1-2	m	urban	low
2	2-7	m	rural	high
3	>7	f	rural	low
4	1-2	f	rural	high
5	>7	m	rural	high
6	1-2	m	rural	high
7	2-7	f	urban	low
8	2-7	m	urban	low

1) ما هي المعالجات الأولية الواجب تطبيقها على المعطيات السابقة قبل إجراء عملية التدريب وفق خوارزمية ID3؟ برر إجابتك.
2) ما هو ناتج تطبيق ID3 على المعطيات السابقة؟
(ملاحظة: لا حاجة لإنجاز الحسابات الدقيقة في حال كانت نتيجة تطبيق ID3 أو أحد مراحلها واضحة! في هذه الحالة يكتفى بتبرير الإجابة فقط).
3) تعاني الشجرة الناتجة عن الطلب السابقة من مشكلة التدريب المفرط overfitting، برر ذلك. اقترح آلية مناسبة للتعامل مع هذه المشكلة، وبين ناتج تطبيقها على المجموعة السابقة.
4) ما رأيك بالآلية المعتمدة في تقسيم قيم السمة الأولى time (إلى ثلاث مجالات كما هو موضح في مجموعة المعطيات)؟ برر إجابتك.
5) لدى اختبار الشجرة الناتجة عن خوارزمية التعلم السابقة على 100 حالة موزعة مناصفة بين درجتي الخطر المحتملتين {low, high}، كان من اللافت قدرة النموذج المدرّب على تصنيف كل حالات الخطر المنخفضة {risk=low} بشكل صحيح، في حين صنفت 20 حالة خطر مرتفع {risk=high} بشكل خاطئ. ارسم مصفوفة الخطأ confusion matrix المعبرة عن هذه النتائج، وأحسب قيم Precision، Recall، F1 لاجمالي الحالات المختبرة.. ما رأيك بجودة النموذج المدرب استناداً إلى نتائج هذه الاختبارات.
6) عند تطبيق خوارزمية التعلم السابقة على مجموعة معطيات أكبر للمسألة ذاتها، لوحظ وجود عدد من حالات التكرار والتضارب. بين تأثير كل من هاتين الحالتين على أداء خوارزمية التعلم، وبيّن المعالجات الأولية الواجب تطبيقها.

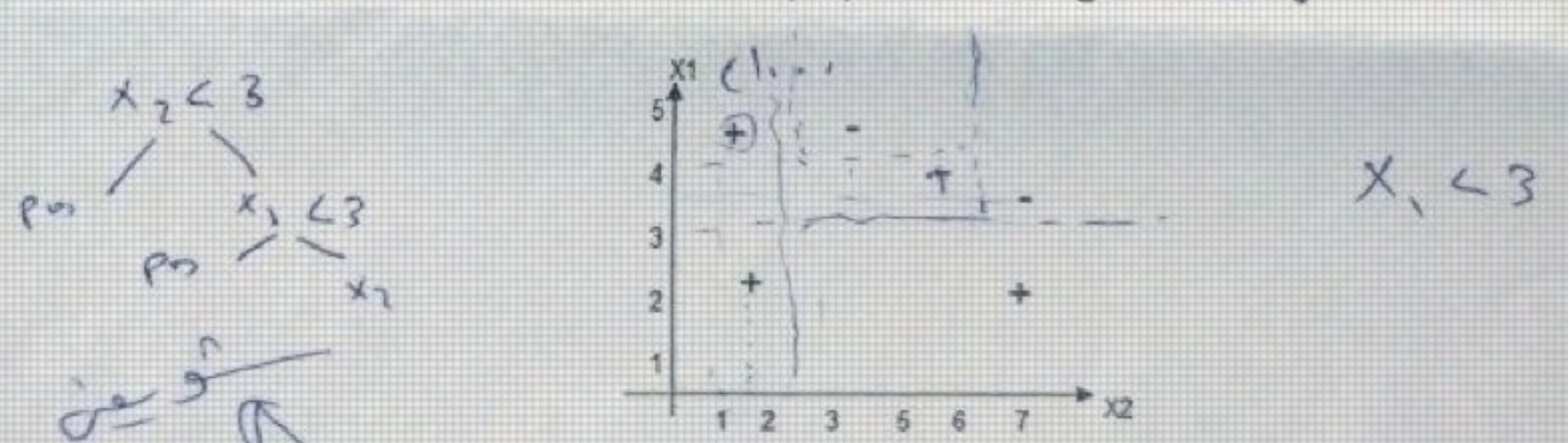
المعهد العالي للعلوم التطبيقية والتكنولوجيا
مذاكرة تعلم الآلة للعام الدراسي 2022-2023
الفصل الأول / السنة الخامسة معلوماتية
المدة: ساعتان / المادة: مفتوحة / يسمح باستعمال الآلات الحاسبة

ملاحظة: في حال لم تتمكن من طرح حل متكامل صحيح لمسألة ما، ضع أفضل إجابة خاطئة توصلت إليها واشرح طريقة محاكمتك للمسألة، ولماذا تظن أن حلّك خاطئ.. الهدف من المذاكرة اختبار طريقة تفكيرك، وليس فقط قدرتك على طرح الأجوبة الصحيحة.

السؤال الأول (25 علامة): أجب عن كلّ من الأسئلة التالية بإيجاز

1) يُعدّ تقسيم المعطيات إلى ثلاث مجموعات من الطرائق شائعة الاستخدام خلال عملية التعلم. ما هي هذه المجموعات، وما دورها في عملية التعلم. وضح ضمن إجابتك دور كلّ منها في تحسين عموميّة النموذج المستنتج.
2) لماذا يُعد أسلوب تقسيم المعطيات وفق مبدأ Leave-one-out Cross-Validation مناسباً في حالة توفر كميّة محدودة من المعطيات.
3) لماذا يعتبر معيار الترابط Correlation بين السمات والخرج المتوقع معياراً غير كافياً لترتيب أهميّة السمات، وبالتالي تخفيض عددها عند الحاجة.
4) ما رأيك بصحة أو خطأ العبارة التالية (برر إجابتك): خوارزميات التعلم مصممة بحيث تعطي أهمية أكبر للسمات الأكثر فائدة. وبالتالي في حال كانت السمات المضافة غير مفيدة فإنها لن تضر بجودة النموذج المتعلم، أي أن اعتماد سمات إضافية يؤدي لتحسين النموذج الناتج عن عملية التعلم أو عدم الإضرار به في أسوأ الأحوال.
5) ما هو دور المعاملات C و Gamma في تحسين عموميّة نموذج SVM الناتج.

السؤال الثاني (15 علامة): لتكن لديك مجموعة المعطيات الممثلة على الشكل التالي، حيث تمثل X_1 و X_2 السمات المستخدمة، في حين يأخذ الخرج قيمتين محتملتين {-,+}



والمطلوب:

1) انقل الشكل السابق إلى ورقة إجابتك وارسم تقديراً لحدود القرار الناتجة وفق طرائق التعلم التالية: أشجار القرار، Linear SVM، KNN، مبرراً طريقة رسمك في كل حالة.
2) أضف مثالاً موجباً في الموقع $(x_1=4, x_2=6)$، أعد الطلب السابق بعد إضافة هذا المثال.

السؤال الثالث (30 علامة): في بعض التطبيقات الطبيّة يكون الحصول على بعض السمات مُكلفاً بدرجة أكبر بكثير من بعضها الآخر (كونها قد تكون ناتجة عن تحاليل أو صور شعاعية باهظة التكلفة مقارنة بسمات أخرى تعكس فحوصات أقل تكلفة). من المفيد في مثل هذه الحالات الحد من الاعتماد على هذه السمات المكلفة ما أمكن ذلك (وهو ما يعني عدم الحاجة لقيم هذه السمات المكلفة إلا في عدد محدود من الحالات)، والمطلوب:

1) يمكن أن يلعب مُعامِل الثقة بنتيجة التنبؤ دوراً هاماً في التعامل مع هذه المسألة، بحيث يكون ناتج التنبؤ مكوناً من الخرج المتوقع إضافة إلى نسبة ثقة النموذج المدرّب بهذه النتيجة (مثلاً: مريض بنسبة ثقة %78

ا- بيّن كيف يمكن تقدير قيمة هذا المعامل اعتماداً على كلٍ من: Naïve Bayes، KNN، SVM، ID3. يمكنك إجراء التعديلات اللازمة على هذه الطرائق في حال الحاجة لذلك.

ب- كيف يمكن الاستفادة من معامل الثقة في تحقيق المطلوب.

2) يمكن التعامل مع المسألة المطلوبة من خلال أخذ تكلفة الحصول على قيمة كل من السمات المدروسة بالاعتبار خلال عملية التعلم ذاتها من خلال التأثير على ترتيب ودور السمات في تحديد شكل النموذج الناتج.

ا- بين لماذا يصعب تطبيق هذا الحل في حالة SVM، بينما يعتبر أكثر قابلية للتطبيق في حالة أشجار القرار.

ب- ما رأيك بقابلية تطبيق هذا المقترح في حالة Naïve Bayes.

ت- اقترح آليّة لتحقيق المطلوب اعتماداً على أشجار القرار.

السؤال الرابع (30 علامة): طُرحت مسألة تحسين خوارزمية ID3 لإيجاد شجرة القرار على مجموعة من الطلاب، وكان الاقتراحان الأكثر ورودا وفق الآتي:

الاقتراح الأول: يجرى التقسيم في كلّ مرة وفق السمة التي تحقق أعلى قيمة ممكنة للمقدار:

$$H(D) - Max(H(D_0), H(D_1))$$

الاقتراح الثاني: يجرى التقسيم في كلّ مرة وفق السمة التي تحقق أعلى قيمة ممكنة للمقدار:

$$H(D) - 0.2*H(D_0) - 0.8*H(D_1)$$

حيث تمثل كل من D, D_0, D_1 مجموعات المعطيات الحالية (قبل التقسيم)، ومجموعات المعطيات بعد التقسيم وفق القيم 0 أو 1 للسمة المعالجة، بينما يمثل H الانتروبية الموافقة لكل مجموعة.

والمطلوب:

1) ما هي السمات التي يتوقع أن تكون أقرب لجذر شجرة القرار في كلٍ من الحالتين السابقتين؟

2) بيّن شجرة القرار الناتجة وفق كلّ من المقترحين السابقين لمجموعة المعطيات التالية:

A	B	C	Output
1	0	1	+
0	1	0	+
0	1	1	-
0	0	1	+

3) ما هو ناتج التنبؤ بخرج المثال التالي (1,1,1) وفق كلّ من الشجرتين الناتجتين عن الطلب السابق.

4) اذكر (مع التبرير) تطبيقاً يكون فيه المقترح الثاني أكثر مناسبة مقارنة بالمقترح الأول.

5) احسب كلاً من الصحة Accuracy، الدقة Precision، الاسترجاع Recall، وقياس F1 عند اختبار الشجرة الناتجة عن المقترح الثاني (في الطلب 2) على مجموعة الاختبار التالية:

A	B	C	Output
1	1	1	+
0	0	0	-
1	0	1	-
1	1	0	+

6) اقترح أحد الطلاب صيغة أكثر ديناميكيّة للمقترح الثاني من خلال إضافة المعامل α وفق الشكل التالي:

$$H(D) - (1- \alpha)*H(D_0) - \alpha*H(D_1)$$

ا- ما هي دلالة المعامل α.

ب- اقترح آليّة لاختيار قيمة مناسبة لهذا المقدار.

مع التمنيات بالتوفيق

د. رياض سنبل

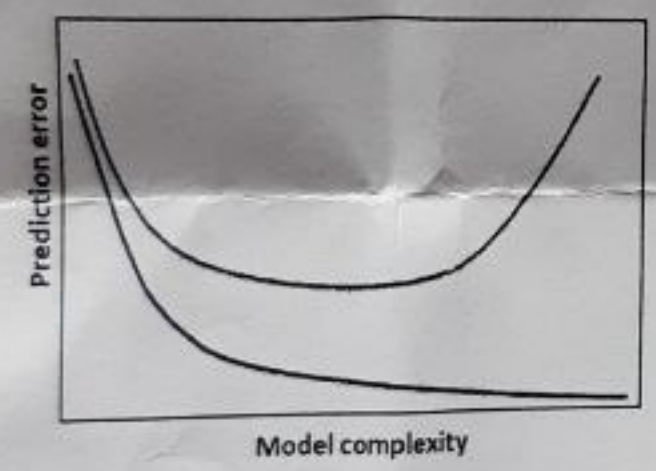
المعهد العالي للعلوم التطبيقية والتكنولوجيا
امتحان تعلم الآلة للعام الدراسي 2022-2023
الفصل الثاني / السنة الرابعة معلوماتية (برمجيات)
المدة: ساعتان / المادة: مغلقة / يسمح باستعمال الآلات الحاسبة

السؤال الأول (15 علامة): أعط رأيك في كلّ من الإدعاءات التالية من جهة صحتها أو احتمالية خطئها مع التبرير:

1) حقق أحد الطلاب خطأ تدريب منخفض في مشروعه، وبناء على ذلك فهو يدعي أن نتيجته جيدة.
2) ادعى فريق عمل أنه حقق نجاحاً كبيراً في مسألة كشف عمليات الاحتيال في عمليات الدفع الإلكتروني بعد تحقيق دقة تصنيف accuracy بنسبة 98 بالمائة في مهمة تصنيف ثنائية على عينة تضمنت 50 مثالًا إيجابيًا و 5000 مثال سلبي.
3) قسّم فريق عمل بياناتهم إلى مجموعتين: تدريب واختبار. استخدام الفريق مجموعة التدريب لاختيار أفضل معاملات محتملة اعتماداً على استراتيجية التدريب Cross-Validation. ومن ثم، جرى بناء نموذج باستخدام هذه المعاملات وكامل مجموعة التدريب، ثم جرى الإبلاغ عن نتائج النموذج الناتج على مجموعة الاختبار.

السؤال الثاني (20 علامة):

1) يوضح الشكل التالي منحنيات التدريب training والتحقق validation لقيم خطأ خوارزمية التعلم مع زيادة تعقيد النموذج. أعد رسم الشكل على ورقة إجابتك وحدد عليه كلاً مما يلي:



أ) أي المنحنيات يرجح أن يكون المنحني المعبر عن خطأ التدريب، وأيها من المرجح أن يعبر عن خطأ التحقق؟
ب) ما هي المناطق التي تعكس تحيزاً bias منخفضاً، وأيها تعكس تحيزاً مرتفعاً؟

2) ما هو التأثير المتوقع على قيم التحيز bias والتباين variance (يزيد، ينقص، لا يتغير) لكل من الحالات التالية، برر إجابتك:
أ) زيادة قيمة K في خوارزمية التعلم KNN.
ب) مضاعفة حجم معطيات التدريب في خوارزمية التعلم Logistic regression.

السؤال الثالث (20 علامة): من أجل كلّ من الثنائيات التالية A, B، اذكر بإيجاز فرقاً جوهرياً يُميّز A عن B:

#	A	B
1)	*Mini-batch gradient descent*	*Stochastic gradient descent*
2)	*Gradient Descent with Momentum*	*Gradient Descent*
3)	*CNN*	*DNN* مهمل
4)	*Bidirectional LSTM*	*LSTM*

السؤال الرابع (30 علامة): لديك مجموعة معطيات من الصور الشعاعية (نفترض أنها باللون الرمادي وأبعادها 192 x 128) موسومة بالصف المقابل لكل منها (سليم، مريض). نريد بناء نظام قادر على تصنيف الصور آلياً. لتحقيق ذلك، اقترح أحد الطلاب استخدام شبكة CNN تليها طبقتين كاملتي الاتصال (fully connected).

1) هل يناسب هذا المقترح حالة مسألتنا المدروسة؟ لماذا؟
2) ما هو الدور المتوقع أن يلعبه كل جزء من أجزاء الشبكة (CNN والطبقات كاملة الارتباط)؟

3) اقترح الطالب اعتماد المعاملات التالية لجزء CNN، حيث يوضح الجدول طبيعة كل طبقة وقياسات النوى المستخدمة لكل منها Kernel size وعددها Kernel، إضافة إلى الخطوة Stride لكل منها)،

Layer	Type	Kernel Size	Stride	Kernel
1	Conv2D	3×3	1	64
2	Conv2D	3×3	1	64
3	Pool	2×2	2	-
4	Conv2D	3×3	1	128
5	Conv2D	3×3	1	128
6	Pool	2×2	2	-
7	Conv2D	3×3	1	256
8	Conv2D	3×3	1	256
9	Pool	2×2	2	-

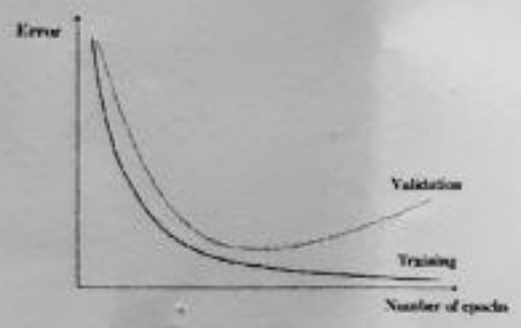
والمطلوب:

أ) ما هو قياس خرج الطبقة الثالثة من جزء CNN من الشبكة. برر إجابتك بالحسابات اللازمة.

ب) هل تعتقد أن خيار الطالب بخصوص البنية التفصيلية لشبكة CNN مبرر؟ في حال كان مبرراً وضح رأيك بالشروحات اللازمة. أما في حال لم يكن مبرراً، وضح كيف كان من الممكن العمل بشكل أكثر منهجية وصولاً لبنية يمكن تبريرها.

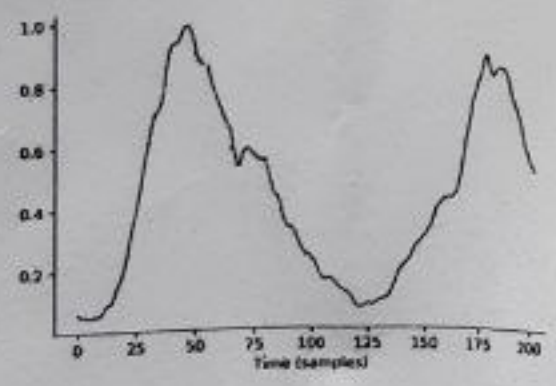
4) اشرح كيف يمكن تدريب جزئي الشبكة؟ ما هو عدد الموسطات الواجب تعلمها في الطبقة الأولى من جزء CNN؟

5) يبيّن الشكل المجاور مخططات لقيم الخطأ على مجموعتي التدريب والتحقق. ما رأيك بسير عملية التدريب؟ اقترح تحسيناً واحداً يساعد على تحسين جودة النموذج المدرب الناتج.



6) بعد إنهاء العمل على هذه المسألة طُلب منك العمل على مسألة تصنيف جديدة تتعلق بحدة المرض اعتماداً على النوعية ذاتها من الصور الشعاعية. اقترح حلاً غير مكلف لإنجاز المطلوب علماً أن عملية التدريب السابقة قد استهلكت تكلفة زمنية وعتادية كبيرة لا يمكن تحملها في المسألة الجديدة.

السؤال الخامس (15 علامة): نتعامل في هذا السؤال مع مسألة التوقع في السلاسل الزمنية، حيث يكون الهدف التوقع والتنبؤ بالقيم المستقبلية في سلسلة زمنية بناءً على سجل زمني للقيم السابقة للمتغير ذاته أو متغيرات ذات صلة به. نكتفي في هذا السؤال بحالة قيم سابقة للمتغير ذاته.



يطلب توضيح النقاط التالية من أجل كل من الطلبات التالية: دخل وخرج وبنية الشبكة ودلالة مكوناتها، وطريقة تجهيز معطيات التدريب والاختبار اعتماداً على قيم السلسلة الزمنية.

1) صمم حلاً يعتمد على شبكة عصبونية بسيطة.
2) صمم حلاً يعتمد على شبكة RNN.

مع التمنيات بالتوفيق

د. رياض سنبل

المعهد العالي للعلوم التطبيقية والتكنولوجيا
مذاكرة تعلم الآلة للعام الدراسي 2022-2023
الفصل الثاني / السنة الرابعة معلوماتية (برمجيات)
المدة: ساعتان / المادة: مفتوحة / يسمح باستعمال الآلات الحاسبة

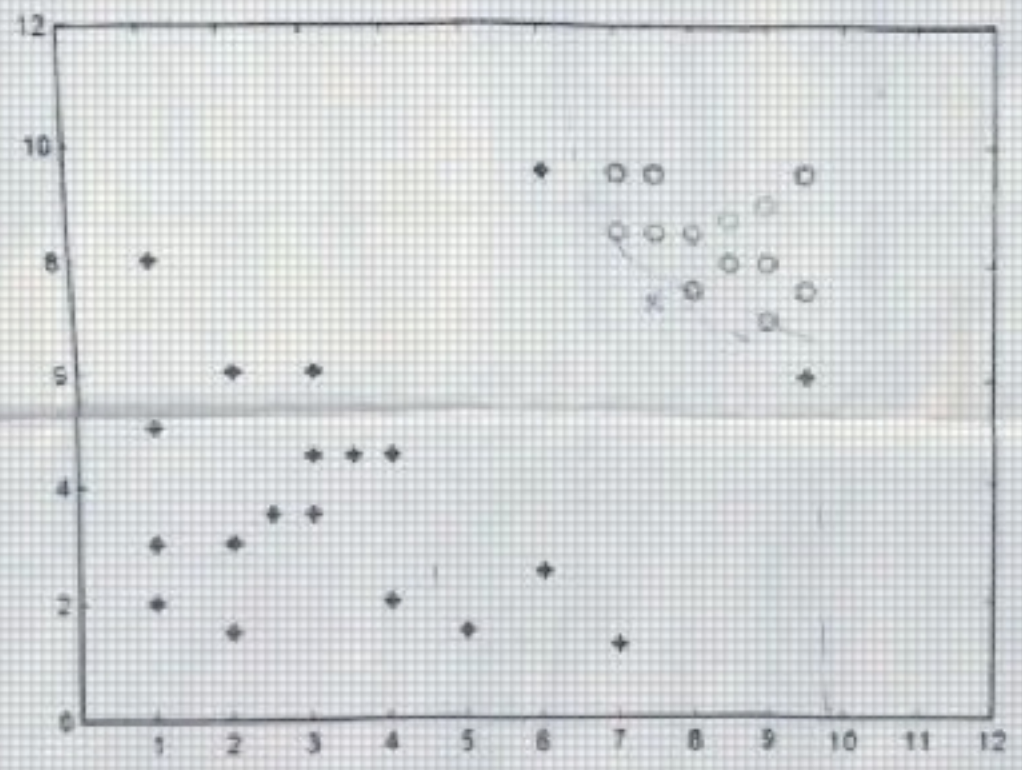
السؤال الأول (20 علامة): نناقش في هذا السؤال المقادير التالية (في سياق تعلم الآلة):

(a) Bias, (b) Variance, (c) Precision,
(d) Out of Sample Error, (e) Irreducible Error

أجب عن الأسئلة التالية من أجل كلّ من هذه المقادير:

1) عمّ يعبر هذا المقدار؟
2) ما دلالة القيم المنخفضة والقيم المرتفعة لهذا المقدار؟
3) أيهما (القيم المنخفضة أم القيم المرتفعة لهذا المقدار) ينعكس إيجاباً على جودة النموذج المدرب؟

السؤال الثاني (30 علامة): نهتم في هذا السؤال بتطبيق خوارزمية SVM للتعلم مع نواة كثير حدود من الدرجة الثانية اعتماداً على مجموعة المعطيات الموضحة في الشكل المبين أدناه. والمطلوب:



1) ارسم حدود القرار الناتجة في حال اعتماد قيمة كبيرة للمعامل C (افترض $C \to \infty$) عند تطبيق خوارزمية SVM مع النواة المذكورة أعلاه (كثير حدود من الدرجة الثانية)، وبين أشعة الدعم Support Vector الموافقة.
2) أعد السؤال السابق بحال كانت قيمة C مساوية للصفر.
3) بفرض أننا نعلم أن مجموعة المعطيات المستخدمة تتضمن نسبة خطأ بمعدل 5%. أي النموذجين السابقين تختار لتحقيق نتائج أفضل عند التطبيق العملي. برر إجابتك.
4) أضف مثالين لمجموعة المعطيات وفق التالي: المثال الأول لا يؤثر على حدود القرار الناتجة عن كل من الطلبين السابقين، المثال الثاني يؤثر بشكل كبير على حدود القرار الناتجة.
5) لنفترض أننا بعد تطبيق خوارزمية SVM، قمنا بحذف الأمثلة المعبرة عن أشعة الدعم Support Vector ثم أعدنا عملية التدريب. ما التغير المتوقع على قيمة الهامش الأعظمي Maximum Margin (تزيد، تنقص، تحافظ على قيمتها). برر إجابتك.

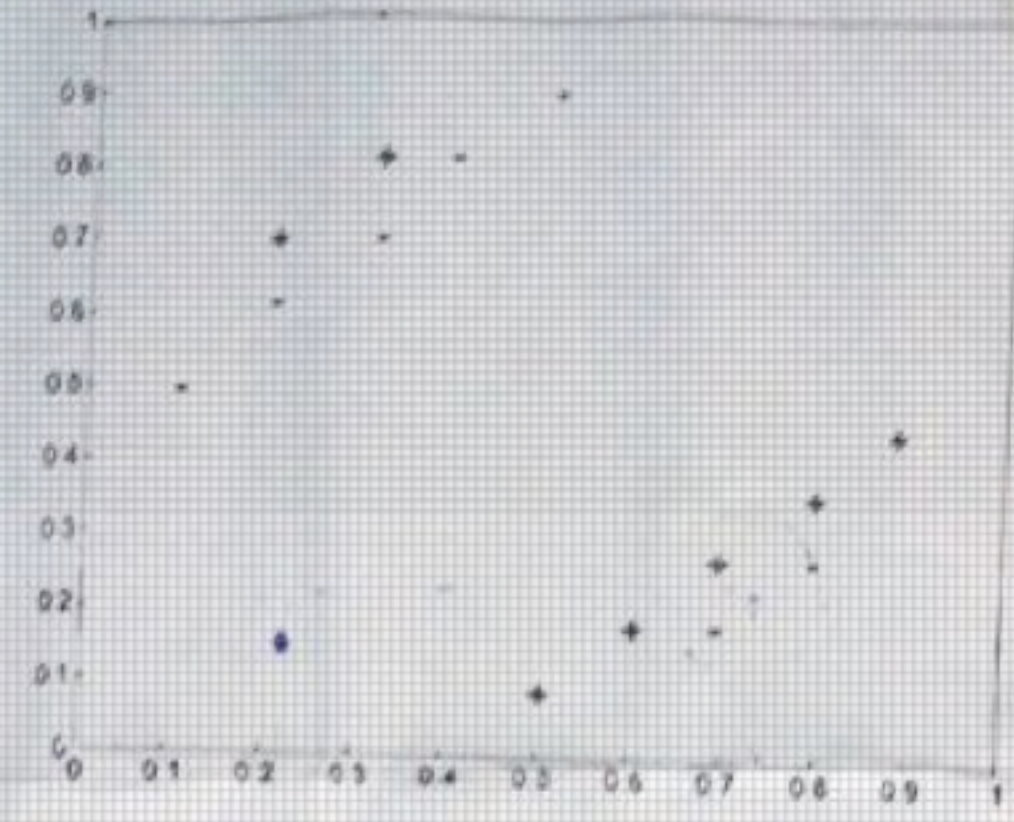
السؤال الثالث (25 علامة): نهتم في هذا السؤال بتطبيق خوارزمية KNN على مجموعة المعطيات المبينة في الشكل أدناه باعتماد المسافة الإقليدية. والمطلوب:

1) ماهي قيمة K التي تحقق أقل خطأ ممكن على عينة التدريب؟ برر إجابتك.
2) لماذا يمكن أن تتأثر جودة النموذج سلباً عند زيادة قيمة K بعد حد معين (وضح هذا الحد من أجل مجموعة المعطيات السابقة)، ولماذا تعد القيمة الصغيرة لـ K بعد حد معين غير موغوبة أيضاً (وضح هذا الحد من أجل مجموعة المعطيات السابقة)؟

3) بفرض جرى اعتماد استراتيجية LOO Cross Validation لاختبار خوارزمية KNN على العينة السابقة. برأيك ماهي قيمة K المثلى؟ برر إجابتك.

4) ارسم حدود القرار الناتجة عن تطبيق KNN مع قيمة K=1 على مجموعة المعطيات السابقة.

5) ما هي دقة النموذج الناتج عن تطبيق استراتيجية LOO Cross Validation مع القيمة K=1، وتابع المسافة Cosine.



السؤال الرابع (25 علامة): لتكن لديك مجموعة المعطيات المبينة في الشكل المجاور، والمكونة من ثلاث سمات A و B و C، إضافة إلى الخرج y.

A	B	C	y
0	0	1	0
0	1	0	0
1	1	0	0
0	0	1	1
1	1	1	1
1	0	0	1
1	1	0	1

1) ما هو عدد الموسطات Parameters الواجب تعلمها وفق Naïve Bayes، قارن هذا العدد مع عدد الموسطات الواجب حسابها وفق Bayes Optimal Classifier.

2) اجر عملية التدريب وفق Naïve bayes واحسب الاحتمالات الناتجة عن ذلك.

3) ما هو توقع النموذج المدرب في الطلب السابق على الحالة التالية A=0, B=0, C=1، ادعم إجابتك بالحسابات اللازمة.

4) إذا علمت أن قيم السمات الثلاث المستخدمة في مجموعة المعطيات مستقلة خطيا فيما بينها. هل يمكنك الجزم بعدم إمكانية تفوق أي مصنف آخر وفق أي طريقة من طرائق تعلم الآلة المدروسة في المقرر على هذا المصنف؟ برر إجابتك.

مع التمنيات بالتوفيق

د. رياض سنبل