



# CONCOURS COMMUN INP RAPPORT DE L'ÉPREUVE ÉCRITE INFORMATIQUE Session 2025

## 1/PRÉSENTATION DU SUJET

### Détection d'objets dans le cadre de la conduite autonome

Le sujet se compose de quatre parties indépendantes. Il apparaissait néanmoins plus pertinent de traiter le sujet de manière linéaire, afin de faciliter la compréhension de l'enchaînement des questions et l'utilisation des fonctions au fur et à mesure.

La première partie porte sur la localisation dans l'espace du véhicule, avec l'utilisation des dictionnaires et des listes.

La deuxième partie concerne le traitement de l'image, incluant les modifications de pixels, le dénombrement, le niveau de gris et l'image intégrale.

La troisième partie est consacrée à un algorithme d'optimisation reposant sur la méthode KNN, laquelle est expliquée et progressivement établie, avec des manipulations de vecteurs et normalisation.

Enfin, la dernière partie porte sur la gestion de base de données.

#### Partie I - Localisation dans l'espace

Cette partie porte sur les dictionnaires et la manipulation de listes, dans le but d'aboutir au calcul de distance focale moyenne entre un véhicule et certains obstacles.

Les obstacles doivent être rangés dans un dictionnaire et la manipulation de listes ainsi que de listes de listes doit permettre d'identifier l'obstacle concerné et de calculer la distance focale associée.

Enfin, une moyenne des focales est demandée à l'issue de cette partie.

#### Partie II - Traitement de l'image

Cette partie débute de manière très classique sur le dénombrement et la taille d'une image de pixels. Des calculs simples relatifs à la conversion en niveau de gris sont ensuite demandés. La complexité de ces traitements est évaluée, puis la méthode d'image intégrale est présentée afin d'améliorer ce traitement d'image.

Les questions relatives à l'image intégrale portent principalement sur la manipulation de matrices de pixels de différentes tailles.

Cette partie se conclut par la détection des obstacles à l'aide d'un algorithme de Viola-Jones qui permet d'apprécier la capacité des candidats à prendre du recul sur l'utilisation des fonctions et la gestion de la taille des matrices de pixels associées.

### **Partie III - Optimisation - Méthode KNN**

Cette partie porte sur la méthode KNN, inscrite au programme de la filière TSI, présentée de manière très progressive avec un grand nombre de questions d'analyse et non de questions de codage. Des fonctions simples utiles à la mise en œuvre de cette méthode sont demandées, notamment la recherche de min et de max, la normalisation, la vectorisation, ainsi que la complétion d'une fonction tri fusion.

L'évaluation de l'algorithme et ses performances sont demandées en fin de partie, en particulier à travers l'utilisation de la matrice de confusion.

### **Partie IV - Gestion dans la base de données**

La dernière partie, très classique, sur la gestion des bases de données, propose une requête à expliciter et trois requêtes à écrire.

## **2/REMARQUES GÉNÉRALES**

### **Erreurs les plus fréquentes**

La notion de dictionnaire en Python n'est pas maîtrisée par certains candidats, notamment en ce qui concerne la construction d'un dictionnaire et l'accès aux doublets (clé, info). La syntaxe associée est fréquemment fautive : utilisation des [] au lieu des {}, parcours non licite d'un dictionnaire, etc.

Les candidats doivent faire preuve de beaucoup plus de rigueur sur ce point.

Des difficultés similaires apparaissent concernant les listes, en particulier pour : " for i in range(L) " au lieu de " for i in L ", ainsi que des problèmes de copies/listes liées.

Les bornes des boucles " for " sont parfois mal traitées, avec par exemple l'utilisation de " range(0,n-1) " au lieu de " range(0,n) ".

On relève également que certains candidats tentent d'utiliser des listes en compréhension alors que ce concept n'est pas suffisamment acquis ; il est donc préférable de conserver une syntaxe plus basique et correctement comprise.

Par ailleurs, une attention accrue doit être portée aux types des variables employées, ainsi qu'aux arguments des fonctions, en particulier lors de leur réutilisation.

La syntaxe des fonctionnalités du module " numpy " n'est pas toujours bien maîtrisée, notamment par ceux qui, par facilité, ont eu recours aux fonctionnalités de ce module.

Des erreurs persistent dans des calculs numériques simples, tels que des additions ou des multiplications. Il est également constaté que certains candidats ne parviennent toujours pas à calculer une moyenne ou un min/max.

La notion de distance euclidienne est souvent confondue avec celle de la distance de Manhattan. De même, la matrice de confusion n'est pas connue, et la méthode KNN reste floue pour certains candidats.

Enfin, de très nombreuses erreurs de syntaxe sont relevées sur le SQL ; il convient d'insister sur la rigueur et la bonne maîtrise des requêtes basiques.

### **Remarques sur le sujet, le texte et sa compréhension**

Le sujet est varié, la thématique retenue est particulièrement intéressante et la difficulté est croissante, ce qui le rend globalement accessible aux candidats et parfaitement en adéquation avec le programme de la filière TSI.

## Problèmes ou difficultés rencontrés lors de la correction

Aucun problème particulier n'a été rencontré lors de la correction.

### Bilan

La grande majorité des candidats ont tenté de traiter un nombre conséquent de questions et ne sont restés bloqués sur aucune d'entre elles. Les questions difficiles et très difficiles ont permis de distinguer clairement les niveaux et d'assurer la sélection entre les candidats.

## 3/REMARQUES SPÉCIFIQUES

### Partie I - Localisation dans l'espace

**Q1.** Un nombre significatif de candidats ne maîtrisent pas le type dictionnaire de Python. Trop de problèmes de syntaxe ont été relevés, bien que les correcteurs aient fait preuve de bienveillance lors de la correction.

**Q2.** Les notions d'accès aux doublets (clé, info) d'un dictionnaire ne sont pas acquises par de nombreux candidats. Certains ont récréé un dictionnaire à partir des données fournies dans le sujet. Bien que cela n'ait pas été sanctionné, ce n'était pas la philosophie attendue de la question.

**Q3.** Plus de la moitié des candidats n'ont pas su répondre correctement. La manipulation des clés et des valeurs d'un dictionnaire est très mal comprise pour la majorité.

**Q4.** Des difficultés importantes apparaissent dans le parcours avec la boucle " for ", notamment concernant l'accès au bon indice. Cette question, un peu délicate en raison de la manipulation de listes de quadruplets associée à un calcul de focale n'a pas été comprise par un grand nombre de candidats.

**Q5.** Cette question a été majoritairement bien traitée, le calcul de moyenne étant trivial.

### Partie II - Traitement de l'image

**Q6.** Des confusions persistent entre la notion de pixel et celle d'octet. Le format RVB n'est pas compris par tous les candidats. Des erreurs ont été relevées dans des opérations d'addition et de multiplication.

**Q7.** Question en général bien traitée. Certains oublient toutefois de diviser C par 255, mais elle reste la question la plus abordée du sujet.

**Q8.** Certains candidats ne respectent pas les arguments en entrée de la fonction : l'entrée n'était pas sous la forme (R, V, B). Il convenait de récupérer ces valeurs depuis " pix " puis de travailler avec la fonction précédente.

**Q9.** Des erreurs ont été relevées sur cette question pourtant relativement simple. De nombreuses erreurs dans la manipulation des listes en compréhension. Des erreurs ont été constatées chez ceux qui utilisent incorrectement la syntaxe des fonctions du module " numpy " (non au programme), comme par exemple la fonction " zeros " pour définir un tableau (array) à double entrée.

**Q10.** Beaucoup d'erreurs sur la taille de la matrice ou de la liste de listes attendue en sortie et sur la réutilisation des fonctions précédentes.

**Q11.** Question globalement bien traitée.

**Q12.** La majorité des candidats ne connaissent pas la division entière. Beaucoup de réponses n'apportaient aucun élément pertinent. Il convenait de comprendre l'intérêt de cette opération vis-à-vis de la méthode d'image intégrale et de sections.

**Q13.** L'intervalle des deux boucles est très souvent mal défini.

**Q14.** Un certain nombre de candidats ont traité cette question correctement. Pour les autres, des erreurs ont été relevées dans l'initialisation du tableau ou de la liste de listes et/ou de la récupération de la taille des deux listes ou des deux composantes du tableau.

**Q15.** Un manque de rigueur est constaté dans la gestion des indices et des listes.

**Q16.** Question difficile : les idées et les tentatives d'explication ont été valorisées. Peu de candidats ont réussi à traiter cette question correctement. Certains n'ont pas compris qu'il était nécessaire d'utiliser des boucles " for " (ou équivalent).

### **Partie III - Optimisation - Méthode KNN**

**Q17.** Question en général bien traitée. Toutefois, certains ne savent pas identifier une valeur max et une valeur min et tentent de trier la liste, ce qui n'était pas demandé.

**Q18.** Peu de candidat ont su répondre correctement, la notion de structure de matrice n'étant pas bien comprise.

**Q19.** Une confusion est observée entre la distance euclidienne et la distance de Manhattan.

**Q20.** Question de cours relative au tri fusion, plutôt bien traitée dans l'ensemble.

**Q21.** Quasiment aucun candidat n'a su répondre correctement. Les complexités des tris sont très mal connues.

**Q22.** Cette question a été traitée correctement pas un certain nombre de candidats. Néanmoins, une mauvaise compréhension de l'algorithme KNN est à l'origine de nombreuses erreurs.

**Q23.** La notion de matrice de confusion n'est quasiment pas connue. Certains candidats l'ont confondue avec une matrice " d'adjacence ", " de voisinage " ou " de cohésion ". Très peu ont été en mesure de donner une valeur, même approximative, du coefficient de performance attendu.

### **Partie IV - Gestion dans la base de données**

**Q24.** Question très facile. Toutefois, il était attendu une explication claire de ce que l'on obtient, et non une simple paraphrase.

**Q25.** Des problèmes de syntaxe ont été relevés, en particulier la notion de " GROUP BY ", qui n'est pas bien maîtrisée par de nombreux candidats.

**Q26.** Des erreurs de syntaxe sont également constatées, notamment concernant les notions de jointure, qui ne sont pas toujours acquises.

**Q27.** Seulement quelques réponses correctes ont été apportées pour cette requête, avec des difficultés persistantes sur la syntaxe.

## 4/ CONCLUSION

Le sujet s'est révélé discriminant. De nombreuses questions étaient abordables, comprenant à la fois des questions particulièrement faciles et d'autres plus difficiles. La longueur du sujet a semblé adapté et en cohérence avec le temps imparti.

### **Conseils aux candidats**

Il est vivement recommandé de maîtriser les algorithmes classiques sur les listes, de savoir manipuler des dictionnaires et les tableaux (arrays), ainsi que de travailler soigneusement le cours relatif aux bases de données.